



## O SIGNIFICADO DE “SIGNIFICADO” \*

Hilary Putnam

Tradução de Alexandre Müller Fonseca

Universidade Federal de Santa Maria – RS - Brasil  
amf.ufsm@gmail.com

A Linguagem é a primeira grande área da capacidade cognitiva humana para a qual estamos começando a obter uma descrição que não é exageradamente simplificada. Graças ao trabalho de linguistas transformacionais contemporâneos,<sup>1</sup> uma descrição muito sutil de ao menos algumas linguagens humanas está sendo construída. Algumas características dessas linguagens parecem ser *universais*. Onde tais características revelam-se “específicas a uma espécie” [*species-specific*] – “não explicáveis sob alguns fundamentos gerais de utilidade funcional ou de simplicidade que se aplicariam a sistemas arbitrários que servem às funções da linguagem” – elas podem lançar alguma luz sobre a estrutura da mente. Embora seja extremamente difícil dizer em que medida a estrutura, assim iluminada, revelar-se-á uma estrutura universal da *linguagem*, por oposição a uma estrutura universal de estratégias gerais de aprendizagem inatas.<sup>2</sup> O próprio fato de que essa discussão possa acontecer é testemunho da riqueza e generalidade do material descritivo que os linguistas estão começando a fornecer, e também testemunho da profundidade da análise, à medida que as características que aparecem como candidatas a características da linguagem “específicas a uma espécie” não são, em nenhum sentido, características superficiais ou fenomenológicas da linguagem, mas residem no nível da estrutura profunda.

---

\* Originalmente publicado em K. Gunderson (ed.), *Language, mind and knowledge*, Minnesota Studies in the Philosophy of Science, vol. VII (Minneapolis: University of Minnesota Press, 1975). Esta tradução baseou-se na versão publicada em H. Putnam, *Mind, Language and Reality: Philosophical Papers, Volume 2* (Cambridge: Cambridge University Press, 1975. pp. 215-271). Os direitos para esta tradução foram graciosamente cedidos pelo autor e pela University of Minnesota Press, atual detentora dos direitos autorais. O trabalho de tradução foi em parte tornado possível por uma bolsa de estudos PROBIC, concedida pela FAPERGS. A tradução foi revista por Rogério Passos Severo (UFSM).

<sup>1</sup> As contribuições a essa área são agora numerosas demais para serem listadas: os pioneiros foram, é claro, Zellig Harris e Noam Chomsky.

<sup>2</sup> Para uma discussão dessa questão, ver H. Putnam, “The ‘innateness’ hypothesis and explanatory models in linguistics”, *Synthese* 17 (1967): 12-22; e N. Chomsky, *Problems of knowledge and freedom* (New York: Vintage Books, 1971), especialmente o capítulo 1.

A desvantagem mais séria de toda essa análise, no que concerne a um filósofo, é que ela não diz respeito ao significado das palavras. A análise da estrutura profunda de formas linguísticas fornece-nos uma descrição incomparavelmente mais poderosa da *sintaxe* de linguagens naturais do que jamais tivemos antes. Mas a dimensão da linguagem associada com a palavra “significado” está, apesar da costumeira enxurrada de tentativas heroicas, senão mal orientadas, tão no escuro quanto sempre esteve.

Neste ensaio, quero explorar o porquê disso ser assim. Em minha opinião, a razão pela qual a assim chamada semântica está em uma condição tão mais precária que a teoria sintática é que o conceito *pré-científico* em que a semântica se baseia – o conceito pré-científico de *significado* – está ele mesmo em uma forma bem pior que o conceito pré-científico de *sintaxe*. Como são habituais em filosofia, dúvidas céticas sobre o conceito em nada auxiliam a esclarecer ou melhorar a situação, como tampouco as asserções dogmáticas proferidas por filósofos conservadores segundo as quais está tudo muito bem neste que é o melhor dos mundos possíveis. A razão pela qual o conceito pré-científico de significado encontra-se em condição tão precária não é esclarecida por alguns argumentos gerais céticos ou nominalistas que têm como conclusão que os significados não existem. De fato, o desfecho de nossa discussão será que os significados não existem inteiramente da maneira como tendemos a pensar que existem. Mas tampouco elétrons existem inteiramente do modo como Bohr os pensou. Há toda uma distância do mundo entre essa afirmação e a afirmação de que significados (ou elétrons) “não existem”.

Falarei quase que exclusivamente do significado de palavras em vez de falar sobre o significado de frases, pois penso que o nosso conceito de significado-de-palavras é mais defeituoso do que nosso conceito de significado-de-frases. Mas comentarei brevemente argumentos de filósofos como Donald Davidson, que insistem que o conceito de significado-de-palavras *tem de ser* secundário e que o estudo de significado-de-frases *tem de ser* primário. Uma vez que considero as teorias tradicionais sobre o significado mitos consumados (note-se que o tópico do “significado” é o único tópico discutido na filosofia em que não há literalmente nada senão “teoria” – literalmente nada que possa ser rotulado ou até mesmo ridicularizado como a ‘concepção de senso comum’), precisarei discutir e tentar desenredar vários tópicos sobre os quais a concepção recebida está, a meu ver, errada. O leitor dar-me-á um grande auxílio na tarefa de tentar esclarecer esses assuntos se gentilmente supuser que *nada* está de antemão claro.

## Significado e extensão

Ao menos desde a Idade Média, autores na teoria do significado propuseram-se a descobrir uma ambiguidade no conceito comum de significado, e introduziram um par de termos – *extensão* e *intensão*, ou *Sinn* e *Bedeutung*, ou o que quer que fosse para desambiguar a noção. A *extensão* de um termo, no jargão lógico costumeiro, é simplesmente o conjunto de coisas das quais o termo é verdadeiro. Desse modo, “coelho”, no seu sentido mais comum em português, é verdadeiro de todos e apenas os coelhos, assim a extensão de “coelho” é precisamente o conjunto dos coelhos. Mesmo essa noção – e trata-se da noção *menos* problemática neste tema nebuloso – tem seus problemas, todavia. Afora os problemas que ela herda da noção aparentada de *verdade*, onde o exemplo anterior de “coelho” *em seu sentido mais corriqueiro* ilustra um desses problemas: estritamente falando, não se trata de

um termo, mas de um par ordenado formado por um termo e um “sentido” (ou uma ocasião de uso, ou alguma outra coisa que distinga um termo num sentido do mesmo termo usado em um sentido distinto), que possui uma extensão. Outro problema é este: um “conjunto”, no sentido matemático, é um objeto de tipo “sim ou não”; qualquer objeto dado ou definitivamente pertence a *S* ou definitivamente não pertence a *S*, se *S* é um conjunto. Mas palavras em uma linguagem natural geralmente não são do tipo “sim ou não”: com certeza existem coisas das quais a descrição “árvore” é claramente verdadeira e coisas das quais a descrição “árvore” é claramente falsa, mas há inúmeros casos limítrofes. Pior do que isso, a linha que separa os casos claros dos casos limítrofes é ela própria mal delimitada [*fuzzy*]. Por conseguinte, a idealização envolvida na noção de extensão – a idealização envolvida na suposição de que há tal coisa como o conjunto de coisas das quais o termo “árvore” é verdadeiro – é na verdade bem severa.

Recentemente, alguns matemáticos investigaram a noção de um *conjunto fuzzy* – isto é, de um objeto ao qual outras coisas pertencem ou não pertencem com uma probabilidade dada ou em um grau dado, em vez de pertencerem ao modo “sim ou não”. Se realmente quiséssemos formalizar a noção de extensão tal como ela é aplicada a termos em uma linguagem natural, seria necessário empregar “conjuntos difusos” ou algo similar, em vez de conjuntos no sentido clássico.

O modo padrão de lidar com o problema de uma palavra ter mais de um sentido é tratar cada um de seus sentidos como se fosse uma palavra diferente (ou melhor, tratar a palavra como se ela carregasse consigo subscritos invisíveis: “burro<sub>1</sub>” – animal de certa espécie; “burro<sub>2</sub>” – ignorante; e como se “coelho<sub>1</sub>” e “coelho<sub>2</sub>”, ou o que quer que seja, fossem palavras inteiramente distintas). Isso envolve novamente duas idealizações muito severas (ou melhor, ao menos duas): supor que as palavras possuem vários sentidos discretos, e supor que o repertório completo dos sentidos está fixado de uma vez por todas. Paul Ziff recentemente investigou em que medida ambas as suposições distorcem a situação real na linguagem natural;<sup>3</sup> não obstante, continuaremos a fazer essas idealizações aqui.

Agora, considere os termos compostos “criatura com um coração” e “criatura com rins”. Supondo-se que toda criatura com um coração possui rins e vice-versa, a extensão desses dois termos é exatamente a mesma. Mas eles obviamente diferem quanto ao significado. Supondo-se que há um sentido de “significado” em que significado = extensão, tem de haver outro sentido de “significado” em que o significado de um termo não é sua extensão, mas outra coisa, digamos, o “conceito” associado ao termo. Chamemos essa “outra coisa” de *intensão* do termo. O conceito de uma criatura com um coração é claramente distinto do conceito de uma criatura com rins. Desse modo, os dois termos possuem intensões diferentes. Quando dizemos que possuem “significados” diferentes, significado = intensão.

## Intensão e extensão

Algo como o parágrafo precedente aparece em toda exposição padrão que se faz das noções de “intensão” e “extensão”. Mas ela não é nem um pouco satisfatória. O porquê de não ser satisfatória, num sentido, é o objeto de todo este ensaio. Mas alguns apontamentos podem ser feitos já desde o início: primeiramente, que evidência temos de que “extensão” é um sentido da palavra “significado”? A explicação canônica das noções de “intensão” e “extensão” é parecidíssima com

<sup>3</sup> Isso é discutido por Ziff em *Understanding understanding* (Ithaca: Cornell University Press, 1972), especialmente no capítulo VIII.

esta: “em um sentido, ‘significado’ significa *extensão*, e noutro sentido ‘significado’ significa *significado*”. O fato é que enquanto a noção de extensão é bastante precisa, relativamente à noção lógica fundamental de *verdade* (e sob as idealizações severas observadas acima), a noção de intensão não fica mais precisa do que a noção vaga (e, como veremos, enganadora) de “conceito”. É como se alguém explicasse a noção de “probabilidade” dizendo: “em um sentido, ‘probabilidade’ significa *frequência*, e em outro sentido significa *propensão*”. “Probabilidade” *nunca* significa “frequência”, e “propensão” é, ao menos, tão pouco clara quanto “probabilidade”.

Por menos clara que seja, a doutrina tradicional, segundo a qual a noção de “significado” possui a ambiguidade extensão/intensão, têm certas consequências típicas. A maioria dos filósofos tradicionais pensou os conceitos como algo *mental*. Desse modo, a doutrina segundo a qual o significado de um termo (isto é, o significado “no sentido de intensão”) é um conceito carrega a implicação de que os significados são entidades mentais. Frege e, mais recentemente, Carnap e seus seguidores, contudo, se rebelaram contra esse “psicologismo”, como eles o denominaram. Percebendo que os significados são propriedades *públicas* – que o *mesmo* significado pode ser “apreendido” por mais de uma pessoa e por pessoas em diferentes momentos –, identificaram conceitos (e consequentemente “intensões” ou significados) com entidades abstratas e não com entidades mentais. Contudo, a “apreensão” dessas entidades abstratas continuava sendo um ato psicológico individual. Nenhum desses filósofos duvidou que compreender uma palavra (conhecer sua intensão) consistia apenas em estar num determinado estado psicológico (mais ou menos como saber multiplicar números de cabeça é apenas uma questão de estar num certo estado psicológico muito complexo).

Em segundo lugar, o velho exemplo dos dois termos “criatura com um coração” e “criatura com rins” mostra que dois termos podem ter a mesma extensão e ainda assim diferir em intensão. Mas tomou-se como óbvio que o inverso é impossível: dois termos não podem diferir em extensão e ter a mesma intensão. É interessante que nenhum argumento para essa impossibilidade jamais foi oferecido. Provavelmente isso reflete a tradição dos filósofos antigos e medievais que supuseram que o conceito correspondente a um termo era apenas uma conjunção de predicados, e, portanto, que o conceito correspondente a um termo sempre *tem* de fornecer uma condição necessária e suficiente para que algo caia na extensão do termo.<sup>4</sup> Para filósofos como Carnap, que aceitaram a teoria verificacionista do significado, o conceito correspondente a um termo fornecia (no caso ideal, onde o termo tinha “significado completo”) um *critério* para pertencer à extensão (não apenas no sentido de “condição necessária e suficiente”, mas no sentido forte de *modo de reconhecer* se uma coisa dada cai na sua extensão ou não). Assim, esses

---

<sup>4</sup> Essa tradição cresceu porque o termo cuja análise provocou toda a discussão na filosofia medieval foi o termo “Deus”, e o termo “Deus” era pensado como sendo definido pela conjunção dos termos “Bom”, “Poderoso”, “Onisciente” etc. – as assim chamadas “Perfeições”. Mas havia um problema, pois Deus era supostamente uma Unidade, e a Unidade era pensada como excluindo Sua essência de *qualquer* complexidade – isto é, “Deus” era definido por uma conjunção de termos, mas Deus (sem aspas) não poderia ser o produto lógico de propriedades, nem poderia ser Ele a única coisa exemplificando o produto lógico de duas ou mais propriedades *distintas*, pois mesmo esse tipo altamente abstrato de “complexidade” era tida como sendo incompatível com Sua perfeição de Unidade. Esse é um paradoxo com o qual teólogos judeus, árabes e cristãos debateram-se durante séculos (por exemplo, na doutrina da Negação da Privação em Maimônides e Tomás de Aquino). É surpreendente que teorias de interesse contemporâneo, tais como o conceitualismo e o nominalismo, foram primeiramente propostas como soluções ao problema da predicação no caso de Deus. É também surpreendente que o modelo favorito de definição em toda essa teologia – o modelo da conjunção-de-propriedades – tenha sobrevivido até os dias de hoje, ao menos nas suas consequências, na filosofia da linguagem.

filósofos positivistas estavam perfeitamente contentes em manter a concepção tradicional neste ponto. Desse modo, a teoria do significado veio a repousar sobre duas suposições não contestadas:

(I) Saber o significado de um termo é apenas uma questão de estar em certo estado psicológico (no sentido de “estado psicológico” em que estados de memória e disposições psicológicas são “estados psicológicos”; ninguém pensou que saber o significado de uma palavra fosse apenas um estado contínuo de consciência, é claro).

(II) O significado de um termo (no sentido de “intensão”) determina sua extensão (no sentido de que a mesma intensão implica a mesma extensão).

Argumentarei que essas duas suposições não são conjuntamente satisfeitas por *qualquer* noção, e muito menos qualquer noção de significado. O conceito tradicional de significado é um conceito que repousa sobre uma teoria falsa.

### ‘Estado psicológico’ e solipsismo metodológico

Para mostrarmos isso, precisamos primeiramente esclarecer a noção tradicional de estado psicológico. Em certo sentido, um estado é simplesmente um predicado binário cujos argumentos são um indivíduo e um tempo. Nesse sentido, *ter 1,65 metros de altura*, *ter dor*, *conhecer o alfabeto*, e até mesmo *estar a mil milhas de Paris* são todos estados. (Note-se que o *tempo* é comumente deixado implícito ou “contextual”; a forma completa de uma frase atômica desses predicados seria “*x tem 1,65 metros de altura no tempo t*”, “*x tem dor no tempo t*” etc.) Em ciência, contudo, é costumeiro restringir o termo “estado” a propriedades que são definidas em termos de parâmetros do indivíduo que são fundamentais do ponto de vista da ciência em questão. Assim, *ter 1,65 metros de altura* é um estado (do ponto de vista da física); *ter dor* é um estado (do ponto de vista da psicologia mentalista, ao menos); *conhecer o alfabeto* pode ser um estado (do ponto de vista da psicologia cognitiva), embora seja difícil dizer; mas *estar a mil milhas de Paris* *não* seria naturalmente chamado de *estado*. Em certo sentido, um estado psicológico é simplesmente um estado estudado ou descrito pela psicologia. Nesse sentido, pode ser trivialmente verdadeiro que, digamos, *saber o significado da palavra “água”* é um “estado psicológico” (visto da perspectiva da psicologia cognitiva). Mas esse não é o sentido de estado psicológico que está em questão na suposição (I) acima.

Quando os filósofos tradicionais falaram sobre estados psicológicos (ou estados “mentais”), fizeram uma suposição que podemos chamar de suposição do solipsismo metodológico. Essa suposição é a de que nenhum estado psicológico, propriamente dito, pressupõe a existência de qualquer indivíduo a não ser o sujeito a quem aquele estado é atribuído. (De fato, a suposição era de que nenhum estado psicológico pressupõe sequer a existência do *corpo* do sujeito: se *P* é um estado psicológico propriamente dito, então tem de ser logicamente possível para uma ‘mente desprovida de corpo’ estar em *P*). Essa suposição é bastante explícita em Descartes, mas está implícita em quase toda a psicologia filosófica tradicional. Obviamente, fazer essa suposição é adotar um *programa restritivo* – um programa que deliberadamente limita o escopo e a natureza da psicologia de modo a ajustá-la a certas pré-concepções mentalistas ou, em alguns casos, a uma reconstrução idealista do conhecimento e do mundo. *Quão* restritivo o programa é, contudo, frequentemente passa despercebido. Estados psicológicos comuns ou prosaicos tais como *ter ciúmes*, precisam ser reconstruídos, por exemplo, se a suposição do solipsismo metodológico é retida. Isso porque, em seu uso comum, *x tem ciúmes de*

y implica que y existe, e x tem ciúmes da estima de y por z implica que tanto y quanto z existem (bem como x, é claro). Assim, ter ciúmes e ter ciúmes da estima de alguém por um terceiro qualquer, não são estados psicológicos permitidos pela suposição do solipsismo metodológico. (Chamar-lhes-emos de “estados psicológicos num sentido amplo” e nos referiremos aos estados que são permitidos pelo solipsismo metodológico como “estados psicológicos no sentido estrito”). A reconstrução exigida pelo solipsismo metodológico seria reconstruir o ciúme de modo que eu possa ter ciúme de minhas próprias alucinações, ou das invenções de minha imaginação etc. Somente se supormos que estados psicológicos no sentido estrito têm um grau significativo de fechamento causal (de modo que nos restringirmos a estados psicológicos no sentido estrito facilitará a formulação de leis psicológicas) há razões para engajarmo-nos nessa reconstrução, ou para supor o solipsismo metodológico. Mas três séculos de fracasso da psicologia mentalista são evidência tremenda contra esse procedimento, em minha opinião.

Seja como for, podemos agora enunciar mais precisamente o que afirmamos ao fim da seção precedente. Sejam A e B dois termos quaisquer que diferem em extensão. Pela suposição (II), eles têm de diferir em significado (no sentido de “intensão”). Pela suposição (I), saber o significado de A e saber o significado de B são estados psicológicos no sentido estrito – pois é assim que interpretaremos a suposição (I). Mas esses estados psicológicos têm de determinar a extensão dos termos A e B tanto quanto os significados (“intensões”) o fazem.

Para vermos isso, vamos tentar supor o contrário. Claro que não pode haver dois termos A e B tal que saber o significado de A é o mesmo estado que saber o significado de B, embora A e B tenham extensões diferentes. Pois saber o significado de A não é somente “apreender a intensão” de A, o que quer que isso venha a ser. É também saber que a “intensão” que foi “apreendida” é a intensão de A. (Desse modo, alguém que saiba o significado de “roda” presumivelmente “apreende a intensão” de seu sinônimo alemão Rad; mas se ele não sabe que a intensão em questão é a intensão de Rad, não se pode dizer que “sabe o significado de Rad”). Se A e B são termos distintos, então saber o significado de A é um estado diferente de saber o significado de B, quer os significados de A e B sejam eles próprios iguais, ou diferentes. Mas, pelo mesmo argumento, se  $I_1$  e  $I_2$  são intensões diferentes e A é um termo, então saber que  $I_1$  é o significado de A é um estado psicológico diferente de saber que  $I_2$  é o significado de A. Assim, não pode haver dois mundos possíveis logicamente diferentes,  $L_1$  e  $L_2$ , tais que, digamos, Oscar esteja no mesmo estado psicológico (em sentido estrito) em  $L_1$  e em  $L_2$  (em todos os aspectos), mas em  $L_1$  Oscar compreende A como tendo o significado  $I_1$  e em  $L_2$  Oscar compreende A como tendo o significado  $I_2$ . (Se assim fosse, então em  $L_1$  Oscar estaria no estado psicológico saber que  $I_1$  é o significado de A e em  $L_2$  Oscar estaria no estado psicológico saber que  $I_2$  é o significado de A, e esses são estados psicológicos, em sentido estrito, diferentes e até mesmo – supondo que A tem apenas um significado para Oscar em cada mundo – incompatíveis.)

Em suma, se S é o tipo de estado psicológico que viemos discutindo – um estado psicológico da forma saber que I é o significado de A, onde I é uma “intensão” e A é um termo – então a mesma condição necessária e suficiente para cair na extensão de A “funciona” em todos os mundos logicamente possíveis em que o falante está no estado psicológico S. Pois o estado S determina a intensão I, e pela suposição (II) a intensão equivale a uma condição necessária e suficiente para pertencer à extensão.

Se nossa interpretação da doutrina tradicional da intensão e da extensão faz jus a Frege e a Carnap, então toda questão do psicologismo/platonismo parece um tanto quanto uma tempestade num copo d'água, no que diz respeito à teoria do significado. (Obviamente, trata-se de uma questão muito importante, no que diz respeito à filosofia geral da matemática.) Pois mesmo se os significados são entidades “platônicas” em vez de entidades mentais na concepção de Frege/Carnap, “apreender” tais entidades consiste presumivelmente em um estado psicológico (no sentido estrito). Ademais, o estado psicológico determina univocamente a entidade “platônica”. Assim, se alguém concebe a entidade “platônica” ou o estado psicológico como o “significado”, isso pareceria algo como uma questão de convenção. E tomar o estado psicológico como sendo o significado dificilmente teria a consequência que Frege temia, isto é: que os significados deixariam de ser públicos. Pois estados psicológicos são “públicos” no sentido em que diferentes pessoas (e mesmo pessoas em diferentes épocas) podem estar no *mesmo* estado psicológico. De fato, o argumento de Frege contra o psicologismo é apenas um argumento contra a identificação de conceitos com particulares mentais, não contra entidades mentais em geral.

O caráter “público” dos estados psicológicos implica, em particular, que se Oscar e Elmer entendem uma palavra *A* de forma diferente, então devem estar em estados psicológicos *diferentes*. Pois o estado de *saber a intensão de A como sendo, digamos, I* é o *mesmo* estado, quer Oscar, quer Elmer estejam nele. Assim, dois falantes não podem estar no mesmo estado psicológico em todos os aspectos e compreender o termo *A* de maneira diferente; o estado psicológico do falante determina a intensão de (e conseqüentemente, pela suposição (II), a extensão) de *A*.

É esta última consequência das suposições (I) e (II) conjugadas que afirmamos ser falsa. Afirmamos que é possível para dois falantes estarem exatamente no *mesmo* estado psicológico (no sentido estrito), mesmo que a extensão do termo *A* no idioleto de um seja diferente da extensão do termo *A* no idioleto de outro. Extensão *não* é determinada pelo estado psicológico.

Isso será mostrado em detalhes nas seções posteriores. Se isso está certo, então há dois percursos abertos àquele que busca resgatar ao menos uma das suposições tradicionais: abandonar a ideia de que estado psicológico (no sentido estrito) determina a *intensão*, ou abandonar a ideia de que a intensão determina a extensão. Consideraremos tais alternativas posteriormente.

## Os significados estão na cabeça?

Que o estado psicológico não determina a extensão será agora mostrado com a ajuda de um pouco de ficção científica. Para o propósito dos seguintes exemplos de ficção-científica, suporemos que em algum lugar da galáxia há um planeta que chamaremos de Terra Gêmea. A Terra Gêmea é muito parecida com a Terra; de fato, as pessoas em Terra Gêmea também falam *inglês*. De fato, afora as diferenças que especificaremos em nossos exemplos ficcionais, o leitor pode supor que a Terra Gêmea é *exatamente* como a Terra. Pode até supor que possui um *Doppelgänger* – uma cópia idêntica – em Terra Gêmea, se quiser, embora minhas histórias não dependam disso.

Embora algumas pessoas em Terra Gêmea (digamos, aqueles que se dizem ‘americanos’ e os que se dizem ‘canadenses’ e os que se dizem ‘ingleses’ etc.) falem inglês, não é surpreendente que haja algumas diferenças mínimas que agora

descreveremos entre os dialetos do inglês falados em Terra Gêmea e o inglês padrão. Essas diferenças dependem de algumas peculiaridades da Terra Gêmea.

Uma dessas peculiaridades de Terra Gêmea é que o líquido chamado “água” não é  $H_2O$ , mas um líquido diferente, cuja fórmula química é muito longa e complicada. Abreviarei essa fórmula química simplesmente como XYZ. Suporei que XYZ é indistinguível da água em condições normais de temperatura e pressão. Em particular, tem gosto de água e sacia a sede como água. Além disso, suporei que os oceanos, lagos e mares de Terra Gêmea contêm XYZ e não água, que chove XYZ em Terra Gêmea e não água etc.

Se uma espaçonave vinda da Terra visitasse Terra Gêmea, então a suposição inicial seria que “água” tem o mesmo significado em Terra e Terra Gêmea. Essa suposição seria corrigida quando fosse descoberto que “água” em Terra Gêmea é XYZ, e a nave terráquea reportaria algo mais ou menos como o que segue:

“Em Terra Gêmea, a palavra ‘água’ significa XYZ.”

(É esse tipo de uso da palavra “significa” que explica a doutrina segundo a qual a extensão é um sentido de significado, a propósito. Mas note-se que embora “significa” significa algo como *tem uma extensão*, nesse exemplo *não* se diria

“Em Terra Gêmea o significado da palavra ‘água’ é XYZ.”

a menos que, possivelmente, o fato de que “água é XYZ” fosse conhecido por todo adulto falante do inglês em Terra Gêmea. Podemos explicar isso em termos da teoria do significado que desenvolvemos abaixo; por ora, apenas observaremos que embora o verbo “significa” por vezes significa “tem uma extensão”, a substantivação “significado” *nunca* significa “extensão”.)

Simetricamente, se uma espaçonave da Terra Gêmea alguma vez visitasse a Terra, então a suposição primeira seria que a palavra “água” tem o mesmo significado na Terra e na Terra Gêmea. Essa suposição seria corrigida quando se descobrisse que “água” na Terra é  $H_2O$ , e a espaçonave de um terráqueo-gêmeo reportaria:

“Na Terra<sup>5</sup> a palavra ‘água’ significa  $H_2O$ .”

Note-se que não há problema quanto à extensão do termo “água”. A palavra simplesmente tem dois significados diferentes (como dizemos): no sentido em que é utilizado em Terra Gêmea, o sentido de água<sub>TG</sub>, o que nós chamamos de “água” simplesmente não é água; enquanto que no sentido em que é usado na Terra, o sentido de água<sub>T</sub>, o que os terráqueos-gêmeos chamam de “água” simplesmente não é água. A extensão de “água” no sentido de água<sub>T</sub> é o conjunto de todos os unidades que consistem em moléculas de  $H_2O$ , ou algo do gênero; a extensão de água no sentido de água<sub>TG</sub> é o conjunto de todas as porções que consistem de moléculas de XYZ, ou algo assim.

Agora vamos voltar no tempo a aproximadamente 1750. Nessa época, a química não havia sido desenvolvida nem na Terra nem na Terra Gêmea. Um falante típico do inglês terráqueo não sabia que água consiste de hidrogênio e

<sup>5</sup> Ou melhor, eles reportariam: “Na Terra Gêmea (*o nome em Terra Gêmea para Terra* – H.P.) a palavra ‘água’ significa  $H_2O$ ”.



oxigênio, e um falante típico do inglês terráqueo-gêmeo não sabia que “água” consiste de XYZ. Seja Oscar<sub>1</sub> um desses falantes típicos do inglês terráqueo, e seja Oscar<sub>2</sub> sua contraparte em Terra Gêmea. Pode-se supor que não há crença que Oscar<sub>1</sub> tenha sobre água que Oscar<sub>2</sub> não tenha sobre “água”. Se você quiser, pode até mesmo supor que Oscar<sub>1</sub> e Oscar<sub>2</sub> são réplicas exatas em aparência, sentimentos, pensamentos, monólogos interiores etc. No entanto, a extensão do termo “água” era H<sub>2</sub>O na Terra tanto em 1750 como em 1950; e a extensão do termo “água” era XYZ em Terra Gêmea tanto em 1750 como em 1950. Oscar<sub>1</sub> e Oscar<sub>2</sub> compreendem o termo “água” diferentemente em 1750, *embora estivessem no mesmo estado psicológico*, e embora, dado o estado da ciência da época, tenha levado as comunidades científicas aproximadamente 50 anos para descobrirem que compreendiam o termo “água” diferentemente. Assim, a extensão do termo “água” (e de fato, seu ‘significado’ no uso intuitivo pré-analítico daquele termo) *não* é por si só uma função do estado psicológico do falante.

Todavia, pode-se objetar, por que deveríamos aceitar que o termo “água” tem a mesma extensão em 1750 e em 1950 (em ambas as Terras)? A lógica de termos para espécies naturais como “água” é uma questão complicada, mas o que se segue é um esboço de resposta. Suponha que eu aponte para um copo d’água e diga “este líquido chama-se água” (ou ‘isto é água’, se o marcador “líquido” estiver claro no contexto). Minha “definição ostensiva” de água possui os seguintes pressupostos empíricos: a massa de líquido para a qual aponto guarda certa relação de identidade (digamos, *x é o mesmo líquido que y*, ou *x é o mesmo<sub>L</sub> que y*) com a maioria das coisas que eu e outros falantes de minha comunidade linguística em outras ocasiões chamamos de “água”. Se essa pressuposição é falsa porque, digamos, estou sem sabê-lo apontando para um copo de gin e não para um copo d’água, então não pretendo que minha definição ostensiva seja aceita. Desse modo, a definição ostensiva transmite o que pode ser chamado de condição necessária e suficiente anulável: a condição necessária e suficiente para ser água consiste em ter a relação mesmo<sub>L</sub> com a substância que está no copo; mas isso é uma condição necessária e suficiente somente se o pressuposto empírico for satisfeito. Se não é satisfeito, então uma de uma série de, por assim dizer, condições “reservas”, irá ativar-se novamente.

O ponto chave é que a relação mesmo<sub>L</sub> é uma relação *teórica*: se algo é ou não o mesmo líquido que *isto*, pode exigir uma quantidade indeterminada de investigação científica para se determinar. Ademais, mesmo se uma resposta “definitiva” for obtida, seja por investigação científica seja pela aplicação de algum teste de “senso comum”, a resposta é *anulável*: uma investigação futura poderá reverter até mesmo o exemplo mais “certo”. Desse modo, o fato de que um falante do inglês em 1750 poderia ter chamado XYZ de “água”, enquanto ele ou seus sucessores não teriam chamado XYZ de água em 1800 ou 1850, não significa que o ‘significado’ de “água” mudou nesse intervalo para o falante médio. Em 1750 ou em 1850 ou em 1950, alguém poderia ter apontado para, digamos, o líquido no lago Michigan como um exemplo de “água”. O que mudou foi que em 1750 teríamos erroneamente pensado que XYZ tinha a relação mesmo<sub>L</sub> com o líquido no lago Michigan, enquanto em 1800 ou em 1850 saberíamos que não (estou ignorando o fato de que o líquido no lago Michigan era água apenas num sentido duvidoso em 1950, é claro).

Modifiquemos agora nossa história de ficção científica. Não sei se panelas e frigideiras podem ser feitas de molibdênio; e, caso seja possível fazê-las de molibdênio, não sei se poderiam ser distinguidas facilmente de panelas e frigideiras

de alumínio. (Não sei nada disso, apesar de ter adquirido a palavra “molibdênio”.) Então suporei que panelas e frigideiras de molibdênio *não podem* ser distinguidas de panelas e frigideiras de alumínio a não ser por um especialista. (Para enfatizar esse ponto, repito que, tanto quanto sei, isso poderia ser verdadeiro, e *a fortiori*, tanto quanto sei, poderia ser verdadeiro em virtude de “saber o significado” das palavras *alumínio* e *molibdênio*). Suporemos agora que molibdênio é tão comum em Terra Gêmea quanto o alumínio na Terra, e que alumínio é tão raro em Terra Gêmea como o molibdênio na Terra. Particularmente, suporemos que panelas e frigideiras de ‘alumínio’ são feitas de molibdênio em Terra Gêmea. Por fim, suporemos que as palavras “alumínio” e “molibdênio” estão trocadas em Terra Gêmea: ‘alumínio’ é o nome de *molibdênio* e “molibdênio” é o nome do *alumínio*.

O exemplo compartilha alguns aspectos com o anterior. Se uma espaçonave da Terra visitasse Terra Gêmea, os visitantes da Terra provavelmente não suspeitariam que as panelas e frigideiras de “alumínio” em Terra Gêmea não eram de alumínio, especialmente quando os terráqueos-gêmeos *dissessem* que são. Mas há uma diferença importante entre os dois casos. Um metalúrgico terráqueo poderia dizer muito facilmente que o “alumínio” era molibdênio, e um metalúrgico de Terra Gêmea poderia dizer com igual facilidade que o alumínio era “molibdênio”. (As aspas na frase precedente indicam os usos em Terra Gêmea.) Enquanto que em 1750 ninguém na Terra ou na Terra Gêmea poderia ter distinguido água de “água”, a confusão de alumínio com “alumínio” envolve apenas uma parte das comunidades linguísticas.

O exemplo tem as mesmas implicações que o precedente. Se Oscar<sub>1</sub> e Oscar<sub>2</sub> são falantes *standard* do inglês terráqueo e do inglês de Terra Gêmea respectivamente, e nenhum dos dois é sofisticado em química ou metalurgia, então pode não haver diferença alguma em seus estados psicológicos quando usam a palavra “alumínio”; não obstante, temos de dizer que “alumínio” tem a extensão de *alumínio* no idioleto de Oscar<sub>1</sub> e a extensão de *molibdênio* no idioleto de Oscar<sub>2</sub>. (Também temos de dizer que Oscar<sub>1</sub> e Oscar<sub>2</sub> querem dizer coisas diferentes com “alumínio”, que “alumínio” tem significados diferentes em Terra e em Terra Gêmea etc.). Novamente, vemos que o estado psicológico do falante *não* determina a extensão (*ou* o “significado”, falando pré-analiticamente) da palavra.

Antes de discutir um pouco mais esse exemplo, deixe-me introduzir um exemplo de ficção *não*-científica. Suponha que você é como eu e não consiga diferenciar um olmo de uma faia. Ainda assim, dizemos que a extensão de “olmo” no meu idioleto é a mesma que a extensão de olmo no de qualquer outra pessoa, a saber, o conjunto de todos os olmos, e que o conjunto de todas as faias é a extensão de “faia” em *ambos* os nossos idioletos. Assim, “olmo” no meu idioleto tem uma extensão diferente de “faia” no seu idioleto (como deveria). É realmente plausível que essa diferença em extensão surja a partir de uma diferença em nossos *conceitos*? Meu *conceito* de olmo é exatamente o mesmo que o meu conceito de faia (envergonho-me de confessar). (Isso mostra que a identificação de significado “no sentido de intensão” com *conceito* não pode estar correta, a propósito). Se alguém tentar heroicamente sustentar que a diferença entre a extensão de “olmo” e a extensão de “faia” no *meu* idioleto é explicada por uma diferença em meu estado psicológico, então podemos sempre refutá-lo construindo um exemplo do tipo “Terra Gêmea” – basta trocar as palavras “olmo” e “faia” em Terra Gêmea (do modo como “alumínio” e “molibdênio” foram trocadas no exemplo anterior). Ademais, suponho que tenho um *Doppelgänger* em Terra Gêmea que é molécula por molécula “idêntico” a mim (no sentido em que dois nós de gravata podem ser “idênticos”). Se

você é um dualista, então suponha também que meu *Doppelgänger* pensa os mesmos pensamentos verbalizados por mim, tem os mesmos dados dos sentidos, as mesmas disposições etc. É absurdo pensar que *seu* estado psicológico seja minimamente diferente do meu: entretanto, ele “quer dizer” *faia* quando diz “olmo” e *eu* “quero dizer” *olmo* quando digo “olmo”. Corte o bolo como você quiser, os “significados” simplesmente não estão na *cabeça*!

### Uma hipótese sócio-linguística

Os dois últimos exemplos dependem de um fato sobre a linguagem que parece, surpreendentemente, nunca ter sido apontado: há uma *divisão do trabalho linguístico*. Dificilmente poderíamos usar palavras tais como “olmo” e “alumínio” se ninguém tivesse uma maneira de reconhecer olmos e alumínio; mas nem todas as pessoas para as quais a distinção é importante têm de ser capazes de fazê-la. Alteremos o exemplo: considere o *ouro*. O ouro é importante por muitas razões: é um metal precioso, tem valor monetário, tem valor simbólico (é importante para a maioria das pessoas que o anel de ouro de casamento que elas usam realmente seja de ouro e não apenas pareça ouro) etc. Considere nossa comunidade como uma “fábrica”: nessa “fábrica”, algumas pessoas têm a “tarefa” de usar anéis de casamento de ouro, outras têm a “tarefa” de *vender anéis de casamento de ouro*, e ainda há outras pessoas que têm a “tarefa” de *dizer se algo é ou não realmente ouro*. Não é nem um pouco necessário ou eficiente que todas as pessoas que usam um anel de ouro (ou uma abotoadura de ouro etc.), ou que discutem o “padrão-ouro” etc., estejam envolvidas na compra e venda de ouro. Nem é necessário ou eficiente que todas as pessoas que compram e vendem ouro sejam capazes de dizer se algo é ou não realmente ouro em uma sociedade onde essa forma de desonestidade seja incomum (a venda de ouro falso) e em que se possa facilmente consultar um especialista em caso de dúvida. E *certamente* não é necessário ou eficiente que todas as pessoas que têm a oportunidade de comprar ou usar ouro sejam capazes de dizer com alguma confiabilidade se algo é ou não realmente ouro.

Os fatos acima mencionados são apenas exemplos da divisão mundana do trabalho (no sentido lato). Mas engendram uma divisão do trabalho linguístico: todas as pessoas para as quais o ouro é importante por alguma razão, têm de *adquirir* a palavra “ouro”; mas não têm de adquirir o *método de reconhecer* se algo é ou não ouro. Elas podem confiar numa subclasse especial de falantes. As características que são geralmente pensadas como estando presentes em conexão com um nome geral – condições necessárias e suficientes para o pertencimento na extensão, modos de reconhecer se algo está na extensão (“critérios”) etc. – estão todas presentes na comunidade linguística *considerada como um corpo coletivo*; mas esse corpo coletivo divide o “trabalho” de conhecer e empregar essas várias partes do “significado” de “ouro”.

Essa divisão do trabalho linguístico repousa e pressupõe também a divisão do trabalho *não-linguístico*, é claro. Se apenas as pessoas que sabem se um metal é realmente ouro ou não tivessem alguma razão para ter a palavra “ouro” em seus vocabulários, então a palavra “ouro” seria como a palavra “água” era em 1750 com respeito àquela subclasse de falantes, e os outros falantes simplesmente não a adquiririam. E algumas palavras não exibem qualquer divisão de trabalho linguístico: “cadeira”, por exemplo. Mas com o aumento da divisão do trabalho na sociedade e o avanço da ciência, mais e mais palavras começam a exibir esse tipo de divisão de trabalho. “Água”, por exemplo, não exibia isso de modo algum antes do surgimento

da química. Hoje é obviamente necessário a todo falante estar apto a reconhecer água (confiavelmente sob condições normais), e provavelmente todo falante adulto conhece até mesmo a condição necessária e suficiente “água é H<sub>2</sub>O”, mas apenas alguns poucos adultos falantes conseguiriam distinguir água de líquidos que superficialmente parecem água. Em caso de dúvida, outros falantes dependem do julgamento desses falantes “especialistas”. Assim, o modo de reconhecimento que esses especialistas possuem é também, por meio deles, patrimônio do corpo linguístico coletivo, mesmo que não seja possuído por cada membro individual desse corpo, e assim o fato mais pesquisado sobre a água pode tornar-se parte do significado *social* da palavra embora seja desconhecido de quase todos os falantes que adquirirem a palavra.

Parece-me que esse fenômeno da divisão do trabalho linguístico é algo que é muito importante que os sociolinguistas investiguem. Em conexão com isso, gostaria de propor a seguinte hipótese:

HIPÓTESE DA UNIVERSALIDADE DA DIVISÃO DO TRABALHO LINGUÍSTICO: Toda comunidade linguística exemplifica o tipo de divisão do trabalho linguístico recém descrito: isto é, possui ao menos alguns termos cujos “critérios” associados são conhecidos apenas por um subconjunto dos falantes que adquirem aqueles termos, e seu uso pelos demais falantes depende de uma cooperação estruturada entres eles e os falantes dos subconjuntos relevantes.

Seria interessante, em particular, descobrir se povos extremamente primitivos foram exceções a essa hipótese (o que indicaria que a divisão do trabalho linguístico é um produto da evolução social), ou se até mesmo eles a exibem. Nesse último caso, pode-se conjecturar que a divisão do trabalho, inclusive do trabalho linguístico, é um traço fundamental de nossa espécie.

É fácil ver como esse fenômeno explica alguns dos exemplos dados acima em que as suposições (I) e (II) falham. Sempre que um termo está sujeito à divisão do trabalho linguístico, o falante ‘médio’ que o adquire não adquire nada que fixe sua extensão. Em particular, seu estado psicológico individual *certamente* não fixa sua extensão; é apenas o estado sociolinguístico do corpo linguístico coletivo ao qual o falante pertence que fixa a extensão.

Podemos resumir essa discussão indicando que há dois tipos de ferramentas no mundo: há ferramentas como martelo e chave de fenda que podem ser usadas por uma pessoa; e há aquelas ferramentas como o navio a vapor que exigem uma atividade cooperativa de várias pessoas. As palavras foram demasiadamente pensadas a partir do modelo do primeiro tipo de ferramenta.

### Indexicalidade e rigidez<sup>6</sup>

O primeiro de nossos exemplos de ficção científica – “água” na Terra e em Terra Gêmea em 1750 – não envolve divisão do trabalho linguístico, ou ao menos não o envolve do mesmo modo que os exemplos de “alumínio” e “olmo”. Não havia (em nossa história, de todo modo) quaisquer “especialistas” em água na Terra em 1750, nem especialistas em “água” em Terra Gêmea. (O exemplo *pode* ser interpretado como envolvendo a divisão do trabalho *ao longo do tempo*, no entanto.

<sup>6</sup> O essencial desta seção foi apresentada numa série de palestras que ministrei na Universidade de Washington (Summer Institute in Philosophy) em 1968, e em uma palestra na Universidade de Minnesota.

Não desenvolverei esse método de tratar o exemplo aqui). O exemplo *de fato* envolve coisas que são de importância fundamental para a teoria da referência e também para a teoria da verdade necessária, que agora discutiremos.

Há duas maneiras óbvias de dizer a alguém o que a gente quer dizer por termos de espécies naturais como os casos de “água”, “limão” e “tigre”. Pode-se dar a elas uma assim chamada definição ostensiva – “isto (este líquido) é água”; “isto (este animal) é um tigre”; “isto (esta fruta) é um limão”; onde os parênteses destinam-se a indicar que os “marcadores” *líquido*, *animal*, *fruta*, podem tanto ser explícitos como implícitos. Ou pode-se fornecer para elas uma *descrição*. Neste último caso, a descrição tipicamente consiste em um ou mais marcadores conjugados a um *estereótipo*<sup>7</sup> – uma descrição padronizada de características da espécie que são típicas, ou ‘normais’, ou de alguma maneira estereotipadas. As características centrais do estereótipo geralmente são *critérios* – características que, em situações normais, constituem modos de reconhecer se uma coisa pertence à espécie ou, ao menos, condições necessárias (ou condições necessárias probabilísticas) para o pertencimento à espécie. Nem todos os critérios usados pela comunidade linguística como um corpo coletivo estão incluídas no estereótipo, e em alguns casos os estereótipos podem ser bastante fracos. Deste modo (a menos que eu seja um falante muito atípico), o estereótipo de um olmo é apenas o de uma árvore decíduifolia comum. Essas características são de fato condições necessárias para o pertencimento à espécie (digo “necessárias” num sentido frouxo, não penso que “olmos são árvores decíduifolias” seja *analítica*), mas estão muito aquém de constituir um modo de reconhecer olmos. Por outro lado, o estereótipo de um tigre de fato permite que tigres sejam reconhecidos (a menos que sejam albinos, ou que alguma outra circunstância atípica esteja presente), e o estereótipo de um limão geralmente permite que limões sejam reconhecidos. Num caso extremo, o estereótipo pode ser *apenas* o marcador: o estereótipo de molibdênio pode ser *apenas* que o molibdênio é um *metal*. Consideremos ambas essas maneiras de introduzir um termo no vocabulário de alguém.

Suponha que eu aponte para um copo contendo um líquido e diga “isto é água”, a fim de ensinar a alguém a palavra “água”. Já descrevemos alguns dos pressupostos empíricos dessa ação, e o modo pelo qual esse tipo de explicação do significado pode fracassar. Vamos tentar agora esclarecer melhor como isso deve ser entendido.

No que se segue, tomaremos a noção de “mundo possível” como primitiva. Fazemos isso porque sentimos que de várias maneiras a noção faz sentido e é cientificamente importante, mesmo que necessite ser tornada mais precisa. Além disso, suporemos que em ao menos alguns casos é possível falar do mesmo indivíduo como existindo em mais de um mundo possível.<sup>8</sup> Nossa discussão inspira-se fortemente no trabalho de Saul Kripke, embora as conclusões tenham sido obtidas independentemente.

Sejam  $M_1$  e  $M_2$  dois mundos possíveis em que eu existo e em que este copo existe, e em que estou dando uma explicação de um significado apontando para este copo e dizendo “isto é água”. (*Não* suporemos que o *líquido* no copo é o mesmo em ambos os mundos). Suponhamos que em  $M_1$  o copo está cheio de  $H_2O$  e em  $M_2$  o copo está cheio de  $XYZ$ . Suporemos também que  $M_1$  é o mundo *atual* e

<sup>7</sup> Ver meu “Is semantics possible?” (in: H. Putnam, *Mind, language and reality: Philosophical papers*, vol. 2. Cap. 8. Cambridge: Cambridge University Press, 1975).

<sup>8</sup> Essa suposição não é realmente necessária no que se segue. O que é necessário é que a mesma *espécie natural* possa existir em mais de um mundo possível.

que XYZ é a substância tipicamente chamada de ‘água’ no mundo  $M_2$  (de modo que a relação entre falantes do inglês em  $M_1$  e falantes do inglês em  $M_2$  é exatamente a mesma que a relação entre falantes do inglês na Terra e falantes do inglês na Terra Gêmea). Há então duas teorias que se podem ter concernentes ao significado de “água”.

(1) Pode-se sustentar que “água” é *relativa ao mundo*, mas *constante* em significado (isto é, a palavra tem um *significado relativo constante*). Nessa teoria, “água” *significa o mesmo* em  $M_1$  e  $M_2$ ; apenas ocorre que água é  $H_2O$  em  $M_1$  e água é XYZ em  $M_2$ .

(2) Pode-se sustentar que água é  $H_2O$  em todos os mundos (a substância chamada ‘água’ em  $M_2$  não é água), mas “água” não tem o mesmo significado em  $M_1$  e  $M_2$ .

Se o que foi dito anteriormente sobre o caso de Terra Gêmea está correto, então (2) é claramente a teoria correta. Quando digo “isto (este líquido) é água”, o “isto” (este líquido) é, por assim dizer, um “isto” *de re* – isto é, a força de minha explicação é que “água” é o que quer que tenha certa relação de equivalência (a relação que chamamos acima de ‘mesmo<sub>L</sub>’) com uma porção do líquido referida como ‘isto’ *no mundo atual*.

Podemos simbolizar a diferença entre as duas teorias como uma diferença de “escopo” da seguinte maneira. Na teoria (1), o seguinte é verdadeiro:

(1’) (Para todo mundo  $M$ ) (Para todo  $x$  em  $M$ ) ( $x$  é água  $\equiv x$  está na relação mesmo<sub>L</sub> com a entidade referida por “isto” em  $M$ )

enquanto na teoria (2):

(2’) (Para todo mundo  $M$ ) (Para todo  $x$  em  $M$ ) ( $x$  é água  $\equiv x$  está na relação mesmo<sub>L</sub> com a entidade referida por “isto” *no mundo atual*  $M_1$ ).

(Chamo isso de diferença de ‘escopo’ porque em (1’) “a entidade referida por ‘isto’” está dentro do escopo de “Para todo mundo  $M$ ”, como a expressão qualificadora “em  $M$ ” deixa explícito, ao passo que em (2’) “a entidade referida por ‘isto’” significa ‘a entidade referida por “isto” *no mundo atual*’, e tem portanto uma referência *independente* da variável ligada “ $M$ ”.)

Kripke chama um designador de “rígido” (em uma frase dada) se (naquela frase) ele se refere ao mesmo indivíduo em todo mundo possível em que o designador designa. Se estendermos a noção de rigidez aos nomes de substâncias, então podemos expressar a teoria de Kripke e a minha dizendo que o termo “água” é *rígido*.

A rigidez do termo “água” segue-se do fato de que, quando forneço a definição ostensiva “isto (este líquido) é água”, quero dizer (2’) e não (1’).

Podemos dizer também, seguindo Kripke, que quando forneço a definição ostensiva “isto (este líquido) é água”, o demonstrativo “isto” é *rígido*.

O que Kripke foi o primeiro a observar é que essa teoria do significado (ou “uso”, ou o que seja) da palavra “água” (e de outros termos de espécies naturais também) tem consequências surpreendentes para a teoria da verdade necessária.

Para explicar isso, deixe-me introduzir a noção de *relação entre-mundos* [cross-world]. Uma relação de dois termos  $R$  será chamada de *entre-mundos* quando for compreendida de tal modo que sua extensão é um conjunto de pares ordenados de indivíduos *que não estão todos no mesmo mundo possível*. Por

exemplo, é fácil entender a relação de *mesma altura* como uma relação entre-mundos: apenas compreenda-a de tal modo que, por exemplo, se  $x$  é um indivíduo em um mundo  $M_1$  que tem 1,7 metros de altura (em  $M_1$ ) e  $y$  é um indivíduo em  $M_2$  que tem 1,7 metros de altura (em  $M_2$ ), então o par ordenado  $x, y$  pertence à extensão de *mesma altura que*. (Uma vez que um indivíduo pode ter alturas diferentes em diferentes mundos possíveis onde aquele mesmo indivíduo existe, estritamente falando não é o par ordenado  $x, y$  que constitui um elemento da extensão de *mesma altura que*, mas antes o par ordenado  $x$ -no-mundo- $M_1$ ,  $y$ -no-mundo- $M_2$ .)

De modo similar, podemos entender a relação *mesmo líquido que* ( $mesmo_L$ ) como uma relação entre-mundos, entendendo-a de tal modo que um líquido no mundo  $M_1$  que tem as mesmas propriedades físicas importantes (em  $M_1$ ) que um líquido em  $M_2$  possui (em  $M_2$ ) está na relação *mesmo<sub>L</sub>* com este último líquido.

Então, a teoria que estamos apresentando pode ser resumida dizendo-se que uma entidade  $x$ , em um mundo possível arbitrário, é *água* se, e somente se, está na relação *mesmo<sub>L</sub>* (construída como uma relação entre-mundos) com as coisas que nós chamamos de “água” no mundo *atual*.

Suponha, agora, que eu não tenha ainda descoberto quais são as propriedades físicas importantes da água (no mundo atual) – isto é, ainda não sei que água é  $H_2O$ . Posso ter modos de *reconhecer* água que são bem-sucedidos (é claro, posso incorrer num pequeno número de erros que não estarei apto a detectar até um estágio posterior de nosso desenvolvimento científico), mas sem conhecer a microestrutura da água. Se concordo que um líquido com as propriedades superficiais de “água” mas com uma microestrutura diferente *não é realmente água*, então meus modos de reconhecer água (minha “definição operacional”, por assim dizer) não pode ser considerada uma especificação analítica de *o que é ser água*. Ou melhor, uma definição operacional, como a ostensiva, é simplesmente um modo de indicar um padrão – indicar as coisas *no mundo atual* de tal forma que para  $x$  ser água, em *qualquer* mundo,  $x$  tem de estar na relação *mesmo<sub>L</sub>* com os membros *normais* da classe de entidades *locais* que satisfazem a definição operacional. “Água” em Terra Gêmea não é água, mesmo que satisfaça a definição operacional, pois não está na relação *mesmo<sub>L</sub>* com as coisas *locais* que satisfazem a definição operacional, e as coisas locais que satisfazem a definição operacional mas têm uma microestrutura diferente do resto das coisas locais que satisfazem a definição operacional tampouco são água, pois não estão na relação *mesmo<sub>L</sub>* com os exemplos *normais* de “água” local.

Suponha agora que eu descubra a microestrutura da água – que água é  $H_2O$ . Nesse momento, estarei apto a dizer que as coisas em Terra Gêmea que anteriormente *confundi* com água não são realmente água. Do mesmo modo, se você descreve não outro planeta no universo atual, mas outro universo possível onde existam coisas com a fórmula química  $XYZ$  que passam pelo ‘teste operacional’ de *água*, teremos de dizer que essas coisas não são água mas meramente  $XYZ$ . Você não terá descrito um mundo possível onde “água é  $XYZ$ ”, mas tão somente um mundo possível onde há lagos de  $XYZ$ , as pessoas bebem  $XYZ$  (e não água), ou o que seja. De fato, uma vez que descobrimos a natureza da água, nada conta como um mundo possível se água não tiver essa natureza. Uma vez que descobrimos que água (no mundo atual) é  $H_2O$ , *nada conta como um mundo possível se nele água não for  $H_2O$* . Em particular, se uma afirmação “logicamente possível” é aquela que se sustenta em alguns “mundos logicamente possíveis”, *não é logicamente possível que água não seja  $H_2O$* .

Por outro lado, podemos perfeitamente bem imaginar-nos tendo experiências que nos convenceriam (e que tornaria racional crer) que água *não* é H<sub>2</sub>O. Nesse sentido, é concebível que água não seja H<sub>2</sub>O. É concebível, mas não é logicamente possível! Ser concebível não é prova de possibilidade lógica.

Kripke refere-se a afirmações que são racionalmente irrevisáveis (supondo-se que existam) como *epistemicamente necessárias*. Às afirmações que são verdadeiras em todos os mundos possíveis, ele se refere simplesmente como necessárias (ou, por vezes, ‘metafisicamente necessárias’). Nessa terminologia, o ponto em questão pode ser reformulado da seguinte forma: uma afirmação pode ser (metafisicamente) necessária e epistemicamente contingente. A intuição humana não possui acesso privilegiado à necessidade metafísica.

Desde Kant, existiu uma grande divisão entre os filósofos que pensam que todas as verdades necessárias são analíticas e filósofos que pensam que algumas verdades necessárias são sintéticas *a priori*. Mas nenhum desses filósofos pensou que uma verdade (metafisicamente) necessária poderia deixar de ser *a priori*: a tradição kantiana foi tão culpada quanto à empirista de igualar a necessidade epistêmica com a metafísica. Nesse sentido, o desafio de Kripke à doutrina recebida vai bem além da usual oscilação empirismo/kantismo.

Neste artigo, entretanto, nosso interesse está na teoria do significado, e não na teoria da verdade necessária. Pontos intimamente relacionados aos de Kripke foram feitos em termos da noção de *indexicalidade*.<sup>9</sup> Palavras como “agora”, “isto”, “aqui”, há muito foram reconhecidas como sendo *indexicais*, ou *token-reflexive* – isto é, têm uma extensão que varia de contexto para contexto ou de *token* para *token*. Para essas palavras ninguém jamais sugeriu a teoria tradicional de que “a intensão determina a extensão”. Tomando o nosso exemplo de Terra Gêmea: se tenho um *Doppelgänger* em Terra Gêmea, então quando penso “eu estou com dor de cabeça”, ele pensa “eu estou com dor de cabeça”. Mas a extensão do *token* particular de ‘eu’ no seu pensamento verbalizado é ele próprio (ou sua classe unitária, para ser preciso), enquanto a extensão do *token* de “eu” no meu pensamento verbalizado sou eu (ou minha classe unitária, para ser preciso). Assim, a mesma palavra, “eu”, tem extensões diferentes em dois idioletos diferentes, mas disso não se segue que o conceito que tenho de mim mesmo seja de alguma maneira diferente que o conceito que meu *Doppelgänger* tem de si mesmo.

Ora, sustentamos que a indexicalidade estende-se para além de palavras e morfemas *obviamente* indexicais (por exemplo, os tempos verbais). Nossa teoria pode ser resumida dizendo-se que palavras como ‘água’ têm um componente indexical despercebido: “água” é aquilo que guarda certa relação de semelhança com a água *ao redor daqui*. Água em outra época ou em outro lugar, ou mesmo em outro mundo possível, tem de estar na relação mesmo<sub>L</sub> com a *nossa* “água” a fim de ser água. Desse modo, a teoria de que (1) palavras têm “intensões”, que são algo como conceitos associados às palavras pelos falantes; e que (2) a intensão determina a extensão – não podem ser verdadeiras de palavras para espécies naturais como “água”, pela mesma razão que a teoria não pode ser verdadeira de palavras obviamente indexicais como “eu”.

A teoria segundo a qual palavras para espécies naturais como “água” são indexicais deixa em aberto, contudo, se devemos dizer que ‘água’ no dialeto inglês de Terra Gêmea tem o mesmo *significado* que “água” no dialeto terráqueo e uma extensão diferente (que é o que normalmente dizemos de “eu” em diferentes

<sup>9</sup> Estes apontamentos foram feitos em minhas palestras de 1968 na Universidade de Washington e na Universidade de Minnesota.



idióletos), desse modo abandonando a doutrina de que “o significado (intensão) determina a extensão”; ou se devemos, como preferimos fazer acima, dizer que uma diferença em extensão é *ipso facto* uma diferença no significado de palavras para espécies naturais, desse modo abandonando a doutrina de que significados são conceitos, ou na verdade, entidades mentais de *qualquer* tipo.

Deveria estar claro, contudo, que a doutrina de Kripke segundo a qual palavras para espécies naturais são designadores rígidos e a nossa doutrina em que são indexicais constituem duas maneiras de construir o mesmo ponto em questão. Queremos energicamente endossar o que Kripke diz quando escreve:

Suponha que de fato fixamos a referência de um nome por uma descrição. Mesmo se assim procedermos, não tornamos esse nome sinônimo da descrição, mas em vez disso usamos o nome rigidamente para referir-nos ao objeto assim nomeado, mesmo falando de situações contrafactuais onde a coisa nomeada não satisfaria a descrição em questão. Ora, isto é o que penso ser de fato verdadeiro naqueles casos de nomear em que a referência é fixada por descrição. Mas, de fato, também penso, ao contrário da maioria dos pesquisadores recentes, que a referência de nomes é raramente ou quase nunca fixada por meio de descrição. E com isso não quero apenas dizer o que Searle diz: “Não se trata de uma descrição única, mas em vez disso de um aglomerado, uma família de propriedades que fixa a referência”. Quero dizer que propriedades nesse sentido não são utilizadas de modo algum.<sup>10</sup>

## Sejamos realistas

Gostaria agora de contrastar minha concepção com outra que é popular, ao menos entre os estudantes (ela parece surgir espontaneamente). Para essa discussão, tomemos como exemplo a palavra para a espécie natural *ouro*. Não distinguiremos entre “ouro” e as palavras cognatas em grego, latim etc. Nosso foco estará em “ouro” no sentido de ouro no estado sólido. Compreendido isso, sustentamos: “ouro” não mudou sua *extensão* (ou não mudou significativamente) em dois mil anos. Nossos métodos para *identificar* ouro sofisticaram-se incrivelmente. Mas a extensão de *χρυσος* no dialeto grego de Arquimedes é a mesma que a extensão de *ouro* no meu dialeto inglês.

É possível (e suponhamos que esse seja o caso) que, assim como havia peças de metal que *não* podiam ser determinadas como sendo de ouro antes de Arquimedes, assim também havia ou há peças de metal que não podiam ser determinadas como *não* sendo de ouro na época de Arquimedes, mas que podemos facilmente distinguir do ouro com as técnicas modernas. Seja *X* uma peça de metal desse tipo. Claramente, *X* não reside na extensão de “ouro” no inglês padrão; minha concepção é que tampouco estava na extensão de *χρυσος* do grego ático, embora um grego antigo tivesse *confundido X* por ouro (ou melhor, por *χρυσος*).

A concepção alternativa é que “ouro” *significa* o que quer que satisfaça a “definição operacional” *contemporânea* de *ouro*. ‘Ouro’, há cem anos, significava o que quer que satisfizesse a “definição operacional” de *ouro* corrente cem anos atrás; “ouro” agora significa o que quer que satisfaça a definição operacional de *ouro* em uso em 1973; e *χρυσος* significa o que quer que satisfaça a definição operacional de *χρυσος* em uso *então*.

<sup>10</sup> S. Kripke, “Identity and necessity”, p. 157 (in: M. Munitz (ed.), *Identity and individuation*. New York: New York University Press, 1972. pp. 135-164).

Um motivo comum para adotar esse ponto de vista é certo ceticismo sobre a *verdade*. Na concepção que estou defendendo, quando Arquimedes dizia que algo era ouro ( $\chi\rho\upsilon\sigma\varsigma$ ), não estava apenas dizendo que possuía as características superficiais do ouro (em casos excepcionais, algo pode pertencer a uma espécie natural e de fato *não* ter as características superficiais de um membro daquela espécie natural). Ele estava dizendo que tinha as mesmas características gerais da *estrutura oculta* (a mesma “essência”, por assim dizer) que qualquer pedaço normal de ouro local. Arquimedes teria dito que nosso pedaço hipotético de metal *X* era ouro, mas ele estaria *errado*. Mas *quem poderia dizer* que ele estaria errado?

A resposta óbvia é: *nós* (usando a melhor teoria disponível hoje). Para a maioria das pessoas, ou a questão (*quem poderia dizer?*) parece relevante, e nossa resposta não, ou nossa resposta parece relevante e a questão não. Por que isso?

A razão, creio, é que as pessoas tendem a ser ou fortemente antirrealistas ou fortemente realistas em suas intuições. Para alguém que tem intuições fortemente antirrealistas faz pouco sentido dizer que o que cai sob a extensão do termo  $\chi\rho\upsilon\sigma\varsigma$  de Arquimedes deve ser determinado usando *nossa* teoria. Pois o antirrealista não vê nossa teoria e a teoria de Arquimedes como duas descrições aproximadamente corretas de algum domínio fixo de entidades independentes de teorias, e ele tende a ser cético sobre a ideia de “convergência” em ciência – ele não vê nossa teoria como uma descrição *melhor* das *mesmas* entidades que Arquimedes estava descrevendo. Mas se nossa teoria é *apenas* a nossa teoria, então usá-la para decidir se *X* cai sob a extensão de  $\chi\rho\upsilon\sigma\varsigma$  é tão arbitrário quanto usar a teoria neandertal para decidir se *X* cai sob a extensão de  $\chi\rho\upsilon\sigma\varsigma$ . A única teoria cujo uso *não* é arbitrário é aquela à qual o próprio falante subscreve.

O problema é que para um antirrealista convicto, *verdade* não faz sentido senão como uma noção intra-teórica (ver meu artigo “Explanation and reference” para uma discussão desse ponto)<sup>11</sup>. O antirrealista pode usar verdade intra-teoricamente no sentido de uma “teoria da redundância”, mas não tem disponíveis *extra-teoricamente* as noções de verdade e referência. No entanto, a *extensão está atrelada à noção de verdade*. A extensão de um termo é apenas aquilo de que o termo é *verdadeiro*. Em vez de tentar reter a noção de extensão via um operacionalismo tortuoso, o antirrealista deveria rejeitar a noção de extensão, como faz com a noção de verdade (em qualquer sentido extra-teórico). Como Dewey, por exemplo, ele pode voltar a uma noção de “asseribilidade justificada” em vez de verdade (relativizada ao método científico, se ele pensa que há um método científico *fixo*, ou aos melhores métodos disponíveis à época, se ele concorda com Dewey que o método científico evolui). Então ele pode dizer que “*X* é ouro ( $\chi\rho\upsilon\sigma\varsigma$ )” era asserível justificadamente na época de Arquimedes e não é asserível justificadamente hoje (de fato, esta é uma afirmação *mínima*, no sentido de que representa um mínimo a respeito do que o realista e o antirrealista podem concordar); mas a asserção de que *X* estava na extensão de  $\chi\rho\upsilon\sigma\varsigma$  será rejeitada como desprovida de significado, bem como a asserção de que “*X* é ouro ( $\chi\rho\upsilon\sigma\varsigma$ )” era *verdadeira*.

É bem sabido que o operacionalismo restrito não consegue explicar com sucesso o uso efetivo de termos científicos e do senso comum. Versões afrouxadas do operacionalismo, como a versão de Carnap acerca da teoria de Ramsey, concordam com, quando não explicam, os usos científicos efetivos (sobretudo porque as versões afrouxadas concordam com quaisquer usos possíveis!), mas

<sup>11</sup> In: H. Putnam, *Mind, language and reality: Philosophical papers*, vol. 2. Cap. 11. Cambridge: Cambridge University Press, 1975.

tornam a comunicabilidade dos resultados científicos um *milagre*. Está fora de questão que os cientistas usam termos como se os critérios associados não fossem *condições necessárias e suficientes*, mas em vez disso, caracterizações *aproximadamente* corretas de algum mundo de entidades independentes da teoria, e que falam como se as últimas teorias num período maduro da ciência fossem, em geral, descrições *melhores* das *mesmas* entidades às quais as teorias anteriores se referiam. Em minha opinião, a hipótese que isso está *certo* é a única hipótese que pode dar conta da comunicabilidade dos resultados científicos, do fechamento das teorias científicas aceitáveis sob a lógica de primeira ordem, e de muitos outros aspectos do método científico.<sup>12</sup> Mas não é a minha tarefa argumentar isso aqui. Meu ponto é que se estamos usando as noções de verdade e extensão em um modo extrateórico (isto é, considerando essas noções como definidas para afirmações expressas nas linguagens de outras teorias que não a nossa própria), então devemos aceitar a perspectiva realista à qual essas noções pertencem. A dúvida sobre se *nós* podemos dizer que *X* não cai sob a extensão de “ouro” tal como *Jones* o usa, é a *mesma* dúvida que aquela sobre se faz sentido pensar na afirmação de *Jones* de que “*X* é ouro” como verdadeira ou falsa (e não apenas “asserível justificadamente por *Jones* e não asserível justificadamente por *nós*”). Ajustar a noção de verdade, que é essencialmente uma noção realista, aos preconceitos antirrealistas que se possa ter, adotando-se uma teoria indefensável do significado não é progresso algum.

Um segundo motivo para se adotar uma explicação operacionalista extremada é uma antipatia por hipóteses não-verificáveis. À primeira vista, pode parecer que estamos dizendo que “*X* é ouro (*χρυσος*)” era falsa à época de *Arquimedes*, embora *Arquimedes* não pudesse, *em princípio*, saber que era falsa. Mas essa não é exatamente a situação. O fato é que há muitas situações que *nós* podemos descrever (usando a própria teoria que nos diz que *X* não é ouro) em que *X* comportar-se-ia de forma bastante diferente do restante das coisas que *Arquimedes* classificava como ouro. Talvez *X* pudesse ser separado em dois metais distintos quando derretido, ou teria tido propriedades de condutibilidade diferentes, ou teria evaporado em uma temperatura diferente, ou o que quer que seja. Se tivéssemos realizado os experimentos com *Arquimedes* nos observando, ele poderia não ter conhecimento da teoria, mas teria sido capaz de verificar a regularidade empírica que “*X* comporta-se diferentemente do resto das coisas que classifico como *χρυσος* em vários aspectos”. Por fim, ele teria concluído que “*X* não pode ser ouro”.

O ponto é que, mesmo se algo satisfaz os critérios usados em um dado momento para identificar ouro (isto é, reconhecer se algo é ouro), pode comportar-se diferentemente do resto das coisas que satisfaz os critérios em uma ou mais situações. Isso pode não *provar* que não se trata de ouro, mas torna relevante a hipótese de que pode não ser ouro, mesmo na ausência de uma teoria. Se, agora, informássemos *Arquimedes* que ouro possui tal e tal estrutura molecular (exceto para *X*), e que *X* comportou-se de forma diferente porque possui uma estrutura molecular diferente, existe alguma dúvida que ele teria concordado conosco que *X* não é ouro? De qualquer forma, preocupar-se porque algo pode ser *verdadeiro* (em qualquer momento dado) e não pode ser *verificado* (naquele momento) parece-me ridículo. Em qualquer concepção razoável, há certamente coisas que são verdadeiras e que não podem ser verificadas em *nenhum* momento. Por exemplo, suponha que há uma infinidade de estrelas binárias. *Temos de ser capazes de*

<sup>12</sup> Para uma discussão esclarecedora justamente desses pontos, ver R. Boyd, *Realism and scientific epistemology* (manuscrito não-publicado: versões fotocopiadas distribuídas pelo autor, Cornell Dept. of Philosophy).

verificar isso, mesmo em *princípio*? (Ver meus artigos “Logical positivism and philosophy of mind”, “Degree of confirmation’ and inductive logic” e “Probability and confirmation”.<sup>13</sup>)

Até aqui lidamos com razões *metafísicas* para rejeitar nossa concepção. Mas alguém pode discordar de nós sobre os fatos empíricos concernentes às intenções dos falantes. Esse seria o caso se, por exemplo, alguém pensasse que Arquimedes (no *Gedankenexperiment* descrito acima) teria dito: “não importa se *X de fato* comporta-se diferentemente de outros pedaços de ouro; *X* é um pedaço de ouro porque *X* possui tais e tais propriedades e isso é tudo que é preciso para ser ouro”. Embora, na verdade, não podemos estar certos de que palavras para espécies naturais na Grécia antiga tinham as propriedades das palavras correspondentes no inglês dos dias atuais, não pode haver qualquer dúvida séria sobre as propriedades dessas últimas. Se deixamos os preconceitos filosóficos de lado, creio que sabemos perfeitamente bem que nenhuma definição operacional fornece uma condição necessária e suficiente para a aplicação de qualquer palavra desse tipo. Podemos fornecer uma “definição operacional”, ou um aglomerado [*cluster*] de propriedades, ou o que for, mas a intenção é nunca “tornar o nome *sinônimo* da descrição”. Em vez disso, “usamos o nome *rigidamente*” para referir-nos a quaisquer coisas que compartilham a *natureza* que as coisas que satisfazem a descrição normalmente possuem.

### Outros sentidos

O que analisamos até aqui foi o sentido predominante de palavras para espécies naturais (ou melhor, a *extensão* predominante). Mas palavras para espécies naturais tipicamente possuem vários sentidos. (Ziff sugeriu até mesmo que elas possuem um *contínuo* de sentidos.)

Parte disso pode ser explicado com base em nossa teoria. Ser água, por exemplo, é estar na relação mesmo<sub>L</sub> relativamente a certas coisas. Mas o que é a relação mesmo<sub>L</sub>?

*x* está na relação mesmo<sub>L</sub> com *y* apenas se (1) *x* e *y* são ambos líquidos e (2) *x* e *y* concordam em propriedades físicas importantes. O termo “líquido” é ele mesmo um termo de espécie natural que não tentarei analisar aqui. O termo “propriedade” é um termo de espectro amplo que analisamos em artigos anteriores. Aquilo em que quero me focar agora é a noção de *importância*. Importância é uma noção relativa a interesses. Normalmente, as propriedades “importantes” de um líquido ou um sólido etc. são aquelas *estruturalmente* importantes: aquelas que especificam aquilo de que o líquido ou sólido etc., em última análise, é feito – partículas elementares, ou hidrogênio e oxigênio, ou terra, ar, fogo, água, ou o que quer que seja – e como estão arranjadas ou combinadas de modo a produzir as características superficiais. Desse ponto de vista, a característica de uma porção típica de água é consistir de H<sub>2</sub>O. Mas pode ser ou não importante que haja impurezas; desse modo, em um contexto “água” pode significar *água quimicamente pura*, enquanto noutro pode significar o que encontramos no lago Michigan. E um falante por vezes pode referir-se a XYZ como sendo água se está usando-o como água. Novamente, normalmente é importante que água esteja no estado líquido; mas algumas vezes isso não importa, e alguém pode referir-se a uma única molécula de H<sub>2</sub>O como água, ou ao vapor d’água como água (“água no ar”).

<sup>13</sup> Publicados em: H. Putnam, *Philosophical papers, volumes 1 and 2* (Cambridge: Cambridge University Press, 1975).

Mesmo sentidos que são tão distantes que têm de ser considerados como sendo um pouco “desviantes”, podem ostentar uma relação definida com o sentido central. Por exemplo, eu poderia dizer “você viu o limão” referindo-me ao limão de *plástico*. Um caso não tão desviante é este: descobrimos “tigres” em Marte. Isto é, eles parecem-se exatamente com tigres, mas possuem uma química à base de silicone em vez de uma base química de carbono. (Um exemplo notável de evolução paralela!). Os “tigres” marcianos são tigres? Isso depende do contexto.

No caso dessa teoria, como no caso de qualquer teoria que seja ortogonal à maneira de como as pessoas pensaram previamente acerca de algo, mal-entendidos certamente surgirão. Um que já surgiu é o seguinte: um crítico manteve que o sentido *predominante* de, digamos, “limão” é aquele em que qualquer coisa que tenha (um número suficiente de) as características superficiais de limão é um limão. O mesmo crítico sugeriu que ter a estrutura oculta – o código genético – de um limão é necessário para ser um limão apenas quando “limão” é usado como um termo *científico*. Ambas essas alegações parecem-me repousar sobre um mal-entendido, ou, talvez, dois mal-entendidos complementares.

O sentido em que literalmente *qualquer coisa* com as características superficiais de um limão é necessariamente um limão, longe de ser o dominante, é extremamente desviante. Nesse sentido, algo seria um limão se parecesse e tivesse o gosto de um limão, mesmo se tivesse uma química à base de silicone, por exemplo, ou mesmo se um microscópio eletrônico revelasse que é uma *máquina*. (Mesmo se incluirmos crescer “como um limão” nas características superficiais, isso não exclui o limão de silicone, se existem pés de “limão” em Marte. Da mesma forma não exclui sequer o limão-máquina; talvez o pé seja uma máquina também!)

Ao mesmo tempo, o sentido em que para ser um limão algo tem de ter o código genético de um limão *não* é o mesmo que o sentido técnico (se existir um, o que duvido). O sentido técnico, suponho, seria aquele em que “limão” fosse *sinônimo* de uma descrição que *especificasse* o código genético. Mas quando dissemos (para mudar o exemplo) que para ser água algo tem de ser H<sub>2</sub>O, não quisemos dizer, como deixamos claro, que o falante tenha de *saber* isso. É apenas por confundir necessidade metafísica com necessidade epistêmica que alguém pode concluir que se a condição de verdade (metafisicamente necessária) para ser água é ser H<sub>2</sub>O, então “água” tem de ser sinônimo de H<sub>2</sub>O – caso em que se trata certamente de um termo científico. E, semelhantemente, embora o sentido predominante de “limão” seja aquele em que para ser um limão algo tem de ter o código genético de um limão (creio eu), não se segue que “limão” seja sinônimo de uma descrição que especifique o código genético explicitamente ou de outro modo.

O erro de se pensar que há um sentido importante de “limão” (talvez o predominante) em que ter as características superficiais de um limão é ao menos *suficiente* para ser um limão, é mais plausível se dentre as características superficiais incluirmos *ser entrecruzável com limões*. Mas a característica de ser entrecruzável com limões pressupõe a noção de ser um limão. Desse modo, mesmo se conseguirmos obter uma condição suficiente *desse modo*, tomar isso como inconsistente com a caracterização oferecida aqui é uma petição de princípio. Além do mais, a caracterização em termos de *limão* pressupondo-se “características superficiais” (como ser entrecruzável com *limões*) não fornece qualquer condição de verdade que nos permitiria decidir que objetos em outros mundos possíveis (ou que objetos um milhão de anos atrás, ou que objetos a um milhão de anos luz daqui) são limões. (Adicionalmente, não penso que essa caracterização, uma vez que incorre em petição de princípio, esteja *correta*, mesmo como uma condição suficiente.

Penso que alguém poderia inventar casos em que algo que não é um limão seria entrecruzável com limões e parecido com um limão etc.)

Novamente, poder-se-ia tentar excluir o caso do limão-máquina (máquina-limão?) que “cresce” numa árvore-máquina (máquina-árvore?) dizendo que “crescer” não é realmente *crescer*. Isso está certo; mas está certo porque *crescer* é um *verbo* de espécie natural, e precisamente o tipo de explicação que estamos apresentando aplica-se a *ele*.

Outro mal-entendido que deveria ser evitado é o seguinte: tomar a explicação aqui desenvolvida como implicando que os membros da extensão de uma palavra para espécie natural necessariamente *possuem* uma estrutura comum oculta. Poderia ter acontecido que as porções do líquido que chamamos de “água” *não* possuem características físicas comuns importantes *exceto* as superficiais. Nesse caso, a condição necessária e suficiente para ser “água” teria sido a posse de um número suficiente de características superficiais.

Incidentalmente, a última afirmação não implica que água poderia ter deixado de ter uma estrutura oculta (ou que água poderia ter sido outra coisa, que não H<sub>2</sub>O). Quando dizemos que poderia ter *acontecido* que a água não tivesse nenhuma estrutura oculta, o que queremos dizer é que um líquido sem qualquer estrutura oculta (isto é, muitas porções de líquidos diferentes, sem nada em comum *exceto* características superficiais) poderia parecer com água, ter gosto de água, e preencher os lagos etc. que são, na realidade, cheios de água. Em suma, poderíamos ter estado na mesma situação epistemológica com relação a um líquido sem nenhuma estrutura oculta como a que estávamos de fato no que diz respeito à água há certo tempo. Compare isso com o que Kripke diz sobre o “púlpito feito de gelo”.<sup>14</sup>

Há, de fato, um número quase contínuo de casos. Descobriu-se que algumas doenças, por exemplo, não têm qualquer estrutura oculta (a única coisa que os casos paradigmáticos têm em comum é um aglomerado de sintomas), enquanto outras têm uma estrutura oculta no sentido de uma etiologia (a tuberculose, por exemplo). Por vezes continuamos sem saber; continua havendo uma controvérsia acirrada acerca do caso da esclerose múltipla.

Um caso interessante é o caso do *jade*. Embora os chineses não reconhecessem uma diferença, o termo “jade” aplica-se a dois minerais: jadeíte e nefrite. Quimicamente, há uma diferença marcante. Jadeíte é uma combinação de sódio e alumínio. Nefrite é feita de cálcio, magnésio e ferro. Essas duas microestruturas bastante diferentes produzem a mesma qualidade textural única!

Voltando ao exemplo de Terra Gêmea por um instante, se H<sub>2</sub>O e XYZ fossem ambos abundantes na Terra, então teríamos tido um caso similar à jadeíte/nefrite: teria sido correto dizer que havia *dois tipos de “água”*. E em vez de dizer que “aquele material de Terra Gêmea revelou-se não ser realmente água”, teríamos que dizer que “revelou-se *do tipo XYZ de água*”.

Para resumir: se há uma estrutura oculta, esta geralmente determina o que é preciso para ser membro de uma espécie natural, não apenas no mundo atual, mas em todos os mundos possíveis. Em outras palavras, ela determina o que podemos e não podemos supor contrafactualmente sobre a espécie natural (“poderia a água ter sido toda ela vapor?” sim / “água poderia ter sido XYZ?” não). Mas a água local, ou o que quer que seja, pode ter duas ou mais estruturas ocultas – ou tantas que a

<sup>14</sup> S. Kripke, “Identity and necessity” (in: M. Munitz (ed.), *Identity and individuation*. New York: New York University Press, 1972).

“estrutura oculta” torna-se irrelevante, e as características superficiais tornam-se decisivas.

## Outras palavras

Até aqui, usamos apenas palavras para espécies naturais como exemplos, mas os apontamentos que fizemos aplicam-se também a muitos outros tipos de palavras. Aplicam-se à grande maioria dos substantivos, e a outras partes do discurso também.

Consideremos por um instante os nomes de artefatos – palavras como “lápiz”, “cadeira”, “garrafa” etc. A concepção tradicional é que essas palavras são certamente definidas por conjunções, ou possivelmente por aglomerados de propriedades. Qualquer coisa com todas as propriedades da conjunção (ou um número suficientemente grande das propriedades do aglomerado, no modelo do aglomerado) é necessariamente *lápiz*, *cadeira*, *garrafa* ou o que quer que seja. Além disso, algumas propriedades do aglomerado (no modelo do aglomerado) são normalmente tidas como sendo *necessárias* (no modelo de conjunção de propriedades, *todas* as propriedades da conjunção são necessárias). *Ser um artefato* é supostamente necessário, e pertencer a um tipo com certo propósito padrão – por exemplo, “lápiz são artefatos”, e “lápiz são tipicamente destinados a servirem para a escrita” são supostamente necessários. Por fim, tal tipo de necessidade é considerado uma necessidade *epistêmica* – de fato, analiticidade.

Vamos mais uma vez fazer um pouco de ficção científica. Dessa vez, usaremos um exemplo criado por Rogers Albritton. Imagine que um dia descobrimos que *lápiz são organismos*. Cortamo-los e os examinamos sob um microscópio eletrônico e vemos os traços quase invisíveis de nervos e outros órgãos. Ficamos espiando e os vemos desovando e seus filhotes crescendo até tornarem-se lápis adultos. Descobrimos que esses organismos não estão imitando outros lápis (artefatos) – não há e nunca houve qualquer lápis exceto esses organismos. É estranho, com certeza que haja *letras* em muitos desses organismos – por exemplo, BONDED *Grants* DELUXE made in U.S.A. No. 2. – talvez sejam organismos inteligentes, e essa seja sua forma de camuflagem. (Também temos que explicar por que ninguém jamais tentou fabricar lápis etc., mas isto é claramente um mundo possível, em algum sentido.)

Se isso é concebível, e eu concordo com Albritton que é, então, é epistemicamente possível que *poderíamos descobrir que lápis são organismos*. Segue-se que *lápiz são artefatos* não é epistemicamente necessário em um sentido forte e, *a fortiori*, não é analítico.

Sejamos cautelosos, contudo. Mostramos que há um mundo possível em que lápis são organismos? Penso que não. O que mostramos foi que existe um mundo possível, onde certos organismos são *contrapartes epistêmicas* de lápis (a expressão é de Kripke). Retornando ao experimento de Terra Gêmea: imagine desta vez que lápis na Terra são exatamente do modo como pensamos que são, artefatos manufaturados para escrever, enquanto “lápiz” em Terra Gêmea são organismos à la Albritton. Imagine, além disso, que isto é totalmente ignorado pelos terráqueos-gêmeos – eles têm exatamente as mesmas crenças sobre “lápiz” que nós temos sobre lápis. Se descobríssemos isso, não diríamos: “alguns lápis são organismos”. Seria muito mais provável que disséssemos: “as coisas em Terra Gêmea que se passam por lápis não são realmente lápis. São, na realidade, uma espécie de organismo”.

Suponha agora que a situação seja como no exemplo de Albritton tanto na Terra como em Terra Gêmea. Então diríamos “lápís são organismos”. Desse modo, quer os “organismos-lápís” em Terra Gêmea (ou noutro universo possível) sejam realmente *lápís* ou não é uma função de se os lápís *locais* são organismos ou não. Se os lápís locais são exatamente o que pensamos que são, então um mundo possível onde existam organismos-lápís *não* é um mundo possível onde *lápís são organismos*; *não* há mundos possíveis onde lápís sejam organismos neste caso (o que é, é claro, o mundo atual). Que lápís sejam artefatos é necessário no sentido de verdadeiro em todos os mundos possíveis – metafisicamente necessário. Mas não se segue que isso seja epistemicamente necessário.

Segue-se que “lápís” não é *sinônimo* de qualquer descrição – nem mesmo frouxamente sinônima de uma descrição *frouxa*. Quando usamos a palavra “lápís”, pretendemos nos referir a qualquer coisa que tenha a mesma *natureza* que os exemplos normais de lápís locais no mundo atual. “Lápís” é tão *indexical* quanto “água” ou “ouro”.

De certa maneira, o caso de lápís revelando-se serem organismos é complementar ao caso que discutimos há alguns anos (ver meu “It ain’t necessarily so”<sup>15</sup>) de gatos revelando-se robôs (controlados remotamente a partir de Marte). Katz<sup>16</sup> argumenta que descrevemos erroneamente esse caso: que o caso seria mais bem descrito como *revelando-se que não há gatos neste mundo*. Katz admite que poderíamos *dizer* “gatos revelaram-se como não sendo animais, mas robôs”; mas ele argumenta que essa é uma frase semanticamente desviante que é glosada como “as coisas às quais me refiro com ‘gatos’ revelaram-se não ser animais, mas robôs”. A teoria de Katz é um caso de linguística mal feita, entretanto. Primeiramente, a explicação de como podemos *dizer* “Gatos são robôs” é simplesmente uma explicação genérica de como podemos dizer *qualquer coisa*. Mais importante, a teoria de Katz prediz que “Gatos são robôs” é *desviante*, enquanto “Não há gatos no mundo” é não-desviante, na verdade, *standard*, no caso descrito. Ora, não nego que *há* um caso em que “não há (e nunca houve) quaisquer gatos no mundo” seria *standard*: poderíamos (epistemicamente falando) descobrir que temos estado sofrendo de uma alucinação coletiva. (“Gatos” são como elefantes cor-de-rosa.) Mas no caso que descrevi, “Gatos revelaram-se como sendo robôs controlados remotamente a partir de Marte” é certamente não desviante, e “não há gatos no mundo” é altamente desviante.

Incidentalmente, as considerações de Katz não são apenas linguística mal feita, mas também uma reconstrução racional mal feita.

A razão pela qual *não* usamos “gato” como sinônimo de uma descrição é certamente que conhecemos o bastante sobre gatos para saber que têm uma estrutura oculta, e é uma boa metodologia científica usar o nome para referir-nos rigidamente às coisas que possuem aquela estrutura oculta, e não ao que quer que por acaso satisfaça alguma descrição. É claro, se *conhecêssemos* a estrutura oculta, poderíamos formular uma descrição em termos *dela*; mas não a conhecemos neste momento. Nesse sentido, o uso de palavras para espécies naturais reflete um fato importante sobre nossa relação com o mundo: sabemos que existem tipos de coisas com estruturas ocultas em comum, mas não temos ainda o conhecimento para descrever todas essas estruturas ocultas.

<sup>15</sup> Publicado em: H. Putnam, *Philosophical papers*, volume 1 (Cambridge: Cambridge University Press, 1975).

<sup>16</sup> J. Katz, “Logic and language: a defense of intentionalism” (in: K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science* 7, 1975).



A concepção de Katz é mais plausível no caso de “lápiz” do que no caso de “gato”, todavia. Pensamos que *conhecemos* uma condição necessária e suficiente para algo ser um *lápiz*, mesmo que vaga. Assim, é possível tornar “lápiz” sinônimo de uma descrição frouxa. Poderíamos dizer, no caso em que “lápiz revelaram-se organismos” ou que “Lápiz revelaram-se organismos” ou que “Não há lápis no mundo” – isto é, poderíamos usar “lápiz” ou como uma palavra para uma espécie natural ou como uma palavra de “um único critério” [“*one-criterion*” word].<sup>17</sup>

Por outro lado, poderíamos duvidar que *haja* verdadeiramente qualquer palavra de um único critério na linguagem natural, afora contextos estipulativos. Não poderia acontecer de os pediatras não serem médicos, mas espiões marcianos? Responda “sim”, e você abandonou a sinonímia entre “pediatra” e “médico especializado no cuidado de crianças”. Parece que há uma tendência forte de que palavras que são introduzidas como tendo “um único critério” desenvolvam um sentido de “espécie natural”, com toda a concomitante rigidez e indexicalidade. No caso dos nomes de artefatos, esse sentido de espécie natural parece ser o predominante.

(Existe uma piada sobre um paciente que está prestes a receber alta de um manicômio. Os médicos o estão questionando por algum tempo, e ele vem dando respostas perfeitamente sãs. Eles decidem deixá-lo partir e ao final da entrevista, um dos médicos casualmente pergunta “O que você quer ser quando sair?” “Uma chaleira”. A piada não seria inteligível se fosse literalmente inconcebível que uma pessoa pudesse ser uma chaleira.)

Há, contudo, palavras que retêm um caráter de um único critério quase puro. Essas são palavras cujos significados derivam de uma transformação: *caçador* = *aquele que caça*.

A concepção apresentada aqui não apenas aplica-se à maioria dos substantivos, mas também a outras partes do discurso. Verbos como “crescer”, adjetivos como “vermelho” etc., todos têm aspectos indexicais. Por outro lado, algumas palavras sincategoremáticas parecem ter mais fortemente o caráter de um único critério. “Inteiro”, por exemplo, pode ser explicado assim: *O exército cercou a cidade* poderia ser verdadeiro mesmo se a divisão A não tivesse participado. *O exército inteiro cercou a cidade* quer dizer que todas as partes do exército (do tipo relevante, por exemplo, a Divisão A) participaram da ação significada pelo verbo.<sup>18</sup>

## Significado

Vejamos agora onde estamos com respeito à noção de significado. Já vimos que a extensão de um termo não é fixada por um conceito que um falante individual tem em sua cabeça, e isso é verdade tanto porque a extensão é, em geral, determinada *socialmente* – há uma divisão do trabalho linguístico tanto quanto do trabalho “real” – e porque a extensão é, em parte, determinada *indexicalmente*. A extensão de nossos termos depende da própria natureza das coisas particulares que servem como paradigmas,<sup>19</sup> e essa natureza própria não é, em geral, plenamente

<sup>17</sup> A ideia de uma palavra de “um único critério”, e uma teoria da analiticidade baseada nessa noção, aparece no meu “The analytic and synthetic” (in: H. Putnam, *Mind, language and reality: Philosophical papers, vol. 2*. Cambridge: Cambridge University Press, 1975).

<sup>18</sup> Esse exemplo vem de uma análise feita por Anthony Kroch (em sua dissertação de doutorado, 1974, Departamento de Linguística, MIT).

<sup>19</sup> Não tenho em mente a noção de Anthony Flew de ‘paradigma’, em que qualquer paradigma de um *K* é *necessariamente* um *K* (na realidade).

conhecida pelo falante. A teoria semântica tradicional deixa de fora apenas duas contribuições para a determinação da extensão – a contribuição da sociedade e a contribuição do mundo real!

Vimos no início que significado não pode ser identificado com extensão. Contudo, não pode ser identificado com “intensão” tampouco, se intensão é algo como um *conceito* de um falante individual. O que devemos fazer?

Há dois caminhos plausíveis que podemos tomar. Um caminho seria reter a identificação do significado com o conceito e pagar o preço de abrir mão da ideia de que o significado determina a extensão. Se seguíssemos esse caminho, poderíamos dizer que “água” tem o mesmo *significado* em Terra e em Terra Gêmea, mas uma *extensão* diferente. (Não apenas uma extensão *local* diferente, mas uma extensão *global* diferente. O XYZ de Terra Gêmea não é a extensão de instâncias de “água” que profiro, mas é a extensão de instâncias de “água” que meu *Doppelgänger* profere, e isso não apenas porque Terra Gêmea está muito longe de mim, uma vez que as moléculas de H<sub>2</sub>O estão na extensão de instâncias de “água” que profiro não importando quão distantes de mim elas estejam no espaço e no tempo. Ademais, o que posso contrafactualmente supor ser água é diferente do que meu *Doppelgänger* pode contrafactualmente supor que “água” seja.) Enquanto esse é o caminho correto a tomar para uma palavra *absolutamente* indexical como “eu”, parece incorreto para as palavras que viemos discutindo. Considere “olmo” e “faia”, por exemplo. Se elas forem “trocadas” em Terra Gêmea, então certamente *não* diríamos que “olmo” tem o mesmo significado na Terra e em Terra Gêmea, mesmo se o estereótipo de uma faia de meu *Doppelgänger* (ou um “olmo”, como ele o chama) for idêntico ao meu estereótipo de um olmo. Em vez disso, diríamos que “olmo” no idioleto de meu *Doppelgänger* significa *faia*. Por essa razão, parece preferível seguir um caminho diferente e identificar “significado” com um par ordenado (ou possivelmente uma *n-upla* ordenada) de entidades, *uma das quais é a extensão*. (Os outros componentes do, por assim dizer, “vetor de significado”, serão especificados posteriormente.) Fazer isso torna trivialmente verdadeiro que o *significado determina a extensão* (isto é, uma diferença em extensão é *ipso facto* uma diferença em significado), mas abandona totalmente a ideia de que há uma diferença no significado que meu *Doppelgänger* e eu atribuímos a uma palavra, então *tem de* haver alguma diferença em nossos conceitos (ou em nosso estado psicológico). Seguindo esse caminho, podemos dizer que meu *Doppelgänger* e eu *queremos dizer coisas diferentes* quando dizemos “olmo”, mas isso não será uma afirmação sobre nossos estados psicológicos. Tudo que isso quer dizer é que as instâncias da palavra que ele profere têm uma extensão diferente das instâncias da palavra que eu profiro; mas essa diferença em extensão não é um reflexo de quaisquer diferenças em nossas competências linguísticas individuais consideradas isoladamente.

Se isso está correto, e penso que está, então o problema tradicional do significado divide-se em dois. O primeiro problema é o de explicar a *determinação da extensão*. Dado que, em muitos casos, a extensão é determinada socialmente e não individualmente, devido à divisão do trabalho linguístico, creio que esse é propriamente um problema para os sociolinguistas. Resolvê-lo, envolveria explicar em detalhes exatamente como a divisão do trabalho linguístico funciona. A assim chamada “teoria causal da referência”, introduzida por Kripke para os nomes próprios e estendida por nós às palavras para espécies naturais e termos para magnitudes físicas (na seção precedente), pertence a esta área. Pois o fato de que, em muitos contextos, atribuímos às instâncias de um nome que profiro o referente, qualquer que ele seja, atribuímos às instâncias do mesmo nome proferido pela

pessoa de quem adquirir o nome (de tal modo que a referência é transmitida de falante para falante, começando com os falantes que estavam presentes na ‘cerimônia de nomeação’, embora nenhuma *descrição* fixa seja transmitida) é simplesmente um caso especial de cooperação social na determinação da referência.

O outro problema é descrever a *competência individual*. A extensão pode ser determinada socialmente em muitos casos, mas não atribuímos a extensão padrão às instâncias de uma palavra *W* proferida por Jones *não importando como* Jones usa *W*. Jones precisa ter algumas ideias e habilidades particulares em conexão com *W* a fim de desempenhar seu papel na divisão do trabalho linguístico. Uma vez que desistimos da ideia de que a competência individual tem de ser tão forte a ponto de efetivamente determinar a extensão, podemos começar a estudá-la com uma mente mais aberta.

Nesse contexto, é instrutivo observar que substantivos como “tigre” e “água” são muito diferentes de nomes próprios. Pode-se usar o nome próprio “Sanders” corretamente sem se saber nada sobre o referente, exceto que se chama “Sanders” – e mesmo isso pode não ser correto. (“Era uma vez, há muito tempo atrás, mais ou menos na sexta-feira passada, Winnie o Urso, que vivia em uma floresta solitariamente sob o nome de Sanders.”) Mas não se pode usar a palavra “tigre” corretamente, salvo *per accidens*, sem conhecer muito sobre tigres, ou ao menos sobre uma certa concepção de tigres. Nesse sentido, conceitos *de fato* têm muito a ver com significado.

Assim como o estudo do primeiro problema é propriamente um tópico em sociolinguística, o estudo do segundo problema é propriamente um tópico em psicolinguística. A este tópico voltamos agora.

## Estereótipos e comunicação

Suponha que um falante saiba que “tigre” tem um conjunto de objetos físicos como sua extensão, e mais nada. Se possui competência linguística normal em outros aspectos, então poderia usar “tigre” em *algumas* frases: por exemplo, “tigres têm massa”, “tigres ocupam espaço”, “dê-me um tigre”, “isto é um tigre?” etc. Além disso, a extensão *socialmente determinada* de “tigre” nessas frases seria a *standard*, isto é, o conjunto dos tigres. No entanto, não considerariamos esse falante como “conhecedor do significado” da palavra *tigre*. Por que não?

Antes de tentar responder a essa questão, vamos reformulá-la um pouco. Falaremos de alguém como tendo *adquirido* a palavra “tigre” se é capaz de usá-la de tal modo que (1) seu uso passa despercebido (isto é, as pessoas não dizem dele coisas como “ele não sabe o que um tigre é”, “ele não sabe o significado da palavra ‘tigre’” etc.); e (2) sua maneira total de se situar no mundo e sua comunidade linguística é tal que a extensão determinada socialmente da palavra “tigre” em seu idioleto é o conjunto de tigres. A cláusula (1) quer dizer, aproximadamente, que falantes como aquele da hipótese do parágrafo precedente não contam como tendo adquirido a palavra “tigre” (ou outra qualquer). Poderíamos falar deles, em alguns casos, como tendo *adquirido parcialmente* a palavra, mas vamos protelar isso por ora. A cláusula (2) significa que os falantes em Terra Gêmea que têm os mesmos hábitos linguísticos que os nossos, contam como tendo adquirido a palavra “tigre” apenas se a extensão de “tigre” no seu idioleto for um conjunto de tigres. A ideia central das seções precedentes deste artigo é que *não* se segue que a extensão de “tigre” no dialeto (ou idioletos) de Terra Gêmea seja o conjunto dos tigres

meramente porque seus hábitos linguísticos são iguais aos nossos: a natureza dos “tigres” de Terra Gêmea é também relevante. (Se os organismos de Terra Gêmea têm uma química à base de silicone, por exemplo, então seus “tigres” não são realmente tigres, mesmo se parecessem com tigres, embora os hábitos linguísticos dos falantes leigos de Terra Gêmea correspondam exatamente aos dos falantes terráqueos.) Dessa forma, a cláusula (2) quer dizer que nesse caso decidimos dizer que falantes de Terra Gêmea não adquiriram nossa palavra “tigre” (embora tenham adquirido outra palavra com a mesma ortografia e pronúncia).

Nossa razão para introduzir esse modo de falar é que a questão “ele sabe o significado da palavra ‘tigre’?” é tendenciosa em favor da teoria segundo a qual adquirir uma palavra é vir a possuir uma coisa chamada seu “significado”. Identifique essa coisa com um conceito, e voltamos à teoria segundo a qual uma condição suficiente para adquirir uma palavra é associá-la ao conceito correto (ou, de forma mais geral, estar no estado psicológico adequado em relação a ele) – exatamente a teoria que gastamos todo esse tempo refutando. Assim, doravante, “adquiriremos” palavras, em vez de “aprender seu significado”.

Podemos agora reformular a questão com a qual esta seção iniciou. O uso do falante por nós descrito não passa despercebido, embora não seja tal que nos leve a atribuir uma extensão não-*standard* à palavra “tigre” em seu idioleto. Por que não passa despercebido?

Suponha que o nosso falante hipotético aponte para uma bola de neve e pergunte “aquilo é um tigre?”. Claramente, não faz muito sentido falar sobre tigres com *ele*. A comunicação significativa exige que pessoas saibam algo daquilo sobre o que estão falando. Decerto, ouvimos todos os dias pessoas “comunicando-se”, muito embora claramente não saibam nada daquilo sobre o que estão falando. Mas o sentido em que o homem que aponta para uma bola de neve e pergunta “aquilo é um tigre?” não sabe nada sobre tigres está tão além do sentido em que o homem que pensa que o Vancouver vai ganhar a Copa Stanley, ou que a guerra do Vietnã foi travada para ajudar os vietnamitas do sul, não sabe do que ele está falando, que chega a desafiar a mente. O problema das pessoas que pensam que o Vancouver vencerá a Copa Stanley, ou que a guerra do Vietnã foi travada para auxiliar os vietnamitas do sul, é algo que obviamente não pode ser remediado pela adoção de convenções linguísticas. Mas não saber sobre o que se está falando no segundo sentido, mentalmente desafiador, pode ser e é evitado, ou quase isso, por nossas convenções linguísticas. O que sustento é que dos falantes *exige-se* que saibam algo sobre tigres (estereotípicos) de modo a poderem contar como tendo adquirido a palavra “tigre”; algo sobre olmos (ou, em todo caso, sobre o estereótipo correspondente) de modo a poder contar como tendo adquirido a palavra “olmo” etc.

Essa ideia não deveria parecer demasiada surpreendente. Afinal, não permitimos que as pessoas dirijam nas estradas sem antes passar por alguns testes a fim de determinar se têm um nível *mínimo* de competência; e não jantamos com pessoas que não aprenderam a usar o garfo e a faca. A comunidade linguística também tem seus padrões mínimos, com respeito tanto à sintática quanto à semântica.

A natureza do nível de competência mínimo exigido depende fortemente tanto da cultura quanto do tópico em questão, todavia. Em nossa cultura, exige-se dos falantes que saibam como os tigres se parecem (se adquirem a palavra “tigre”, e isso é virtualmente obrigatório); deles não se exige que saibam os detalhes mínimos (tal como o formato das folhas) de como um olmo se parece. Dos falantes do inglês,

a comunidade linguística exige que sejam capazes de distinguir tigres de leopardos; não se exige que sejam capazes de distinguir um olmo de uma faia.

Isso poderia facilmente ter sido diferente. Imagine uma tribo indígena, chamemo-la Cheroquoi, que têm as palavras, digamos, *uhaba'* e *wa'arabi* para olmos e faias respectivamente, e que seja *obrigatório* o conhecimento da diferença. De um Cheroquoi que não conseguisse reconhecer um olmo dir-se-ia que não sabe o que uma *uhaba'* é, que não sabe o significado da palavra *uhaba'* (talvez que não conhece a palavra, ou que não *tem* a palavra); assim como de um falante do inglês que não faz ideia que tigres são listrados dir-se-ia que não sabe o que é um tigre, que não sabe o significado da palavra “tigre” (é claro, se ao menos sabe que tigres são grandes felinos, poderíamos dizer que sabe parte do significado, ou que sabe parcialmente o significado) etc. Desse modo, a tradução de *uhaba'* como “olmo” e *wa'arabi* como “faia” seria, em nossa concepção, apenas *aproximadamente* correta. Nesse sentido, há uma dificuldade real com a tradução radical,<sup>20</sup> mas essa não é a dificuldade abstrata da qual Quine fala.<sup>21</sup>

### O que são os estereótipos

Introduzi a noção de um “estereótipo” em minhas palestras na Universidade de Washington e no Minnesota Center para a Filosofia da Ciência em 1968. O artigo “Is semantics possible?” subsequentemente publicado<sup>22</sup> avança na argumentação, e no presente ensaio quero introduzir novamente a noção e responder algumas questões que foram endereçadas a ela.

Na fala ordinária, um “estereótipo” é uma ideia (frequentemente maliciosa) convencional (que pode ser altamente imprecisa) do que um *X* parece ser, ou como age, ou é. Obviamente, estou lidando com alguns aspectos da fala comum. Não estou interessado em estereótipos maliciosos (salvo onde a linguagem ela própria é maliciosa); mas estou interessado em ideias convencionais, que podem ser imprecisas. Estou sugerindo apenas que uma ideia convencional desse tipo está associada a “tigre”, “ouro” etc. e, além disso, que esse é o único elemento de “verdade” na teoria do “conceito”.

Nessa concepção, de alguém que sabe o que “tigre” quer dizer (ou, como alternativamente decidimos dizer, de alguém que adquiriu a palavra “tigre”) *exige-se* que saiba que tigres *estereotípicos* são listrados. Mais precisamente, há *um* estereótipo de tigres (haver outros) que é exigido pela comunidade linguística como tal; exige-se dela que conheça esse estereótipo, e que saiba (implicitamente) que é obrigatório. Esse estereótipo tem de incluir a característica das listras para que sua aquisição seja considerada bem-sucedida.

O fato de que uma característica (por exemplo, as listras) esteja incluída no estereótipo associado a uma palavra *X*, não significa que seja uma verdade analítica que todos os *Xs* têm essa característica, nem que a maioria dos *Xs* tem essa característica, nem que todos os *Xs* normais têm essa característica, nem que

---

<sup>20</sup> O termo deve-se a Quine (em *Word and object*): significa uma tradução não baseada em pistas ou indicações provenientes de culturas compartilhadas ou cognatas.

<sup>21</sup> Para uma discussão da suposta impossibilidade de uma única tradução radical correta, ver o meu “The refutation of conventionalism” (in: H. Putnam, *Mind, language and reality: Philosophical papers, vol. 2*. Cambridge: Cambridge University Press, 1975).

<sup>22</sup> No capítulo 9 de meu *Mind, language and reality: Philosophical papers, Vol. 2*. Cambridge: Cambridge University Press, 1975.

alguns Xs têm essa característica.<sup>23</sup> Tigres de três pernas e tigres albinos não são entidades logicamente contraditórias. Descobrir que nosso estereótipo tem se baseado em membros não-representativos ou não-normais de uma espécie natural, não é descobrir uma contradição lógica. Se os tigres perdessem suas listras, não deixariam de serem tigres, nem borboletas necessariamente deixariam de serem borboletas se perdessem suas asas.

(Falando estritamente, a situação é mais complicada. É possível dar a uma palavra como “borboleta” um sentido em que borboletas cessariam de serem borboletas se perdessem suas asas – por meio de mutação, digamos. Desse modo, pode-se encontrar *um* sentido de “borboleta” em que “borboletas têm asas” seja analítico. Mas o sentido mais importante do termo, creio eu, é aquele no qual borboletas sem asas continuariam sendo borboletas.)

Nesta altura, o leitor pode se perguntar qual é o valor para a comunidade linguística de se ter estereótipos, se a “informação” contida no estereótipo não é necessariamente correta. Mas isso na verdade não é tão misterioso. A maioria dos estereótipos de fato captam aspectos que os membros paradigmáticos da classe em questão possuem. Mesmo onde os estereótipos erram, o modo pelo qual erram lança luz sobre a contribuição normalmente feita pelos estereótipos à comunicação. O estereótipo de ouro, por exemplo, contém a característica *amarelo* apesar de o ouro quimicamente puro ser quase branco. Mas o ouro que vemos em joalherias é tipicamente amarelo (devido à presença de cobre); assim, a presença dessa característica no estereótipo é até mesmo útil em contextos leigos. O estereótipo associado a *bruxa* é mais equivocado ainda, ao menos se considerado como tendo significação existencial. Acreditar (com significação existencial) que bruxas fazem pactos com Satã, que causam doenças e mortes etc. facilita a comunicação apenas no sentido de facilitar a comunicação interna à teoria das bruxas. Não facilita a comunicação em qualquer situação onde o que é necessário é mais um acordo com o mundo do que um acordo com a teoria dos outros falantes. (Falando estritamente, estou falando do estereótipo como existiu na Nova Inglaterra há 300 anos; hoje, as bruxas não serem *reais* faz parte do estereótipo, e os efeitos nocivos da teoria das bruxas são desse modo neutralizados.) Mas o fato de que nossa linguagem possui *alguns* estereótipos que mais dificultam do que facilitam nosso entendimento com o mundo e entre nós mesmos apenas indica o fato de que não somos seres infalíveis, e como poderíamos ser? O fato é que dificilmente poderíamos nos comunicar com sucesso se a maioria de nossos estereótipos não fossem bastante precisos no que lhes concernem.

### O “significado operacional” dos estereótipos

Uma questão mais traiçoeira é esta: em que medida a noção de estereótipo é “operacionalmente definível”. Aqui é necessário ser extremamente cauteloso. Tentativas nas ciências físicas para especificar literalmente definições operacionais de termos têm notoriamente fracassado; e não há razão para a tentativa ser bem-sucedida em linguística se falhou na física. Por vezes, os argumentos de Quine contra a possibilidade de uma teoria do significado parecem reduzir-se à demanda por definições operacionais em linguística; quando esse é o caso, os argumentos serão ignorados. Mas acontece frequentemente que termos têm de fato definições operacionais não no mundo atual, mas em circunstâncias idealizadas. Fornecer

<sup>23</sup> Argumento em favor disso em “Is semantics possible?” (in: H. Putnam, *Mind, language and reality: Philosophical papers*, vol. 2. Cambridge: Cambridge University Press, 1975).

essas “definições operacionais” tem um valor heurístico, como a idealização frequentemente também tem. É apenas quando tomamos erroneamente a definição operacional como sendo algo mais do que uma idealização conveniente que ela se torna prejudicial. Assim, podemos perguntar: qual é o “significado operacional” da afirmação de que uma palavra tem um estereótipo tal e tal, sem supor que a resposta a essa questão conta como sendo uma explicação teórica do que é ser um estereótipo.

A explicação teórica do que é ser um estereótipo procede em termos da noção de *obrigação linguística* – uma noção que acredito ser fundamental à linguística e que não tentarei explicar aqui. O significado de se dizer que ser listrado é parte do estereótipo (linguístico) de “tigre” é que é *obrigatório* adquirir a informação de que os tigres estereotípicos são listrados para se adquirir “tigre”, no mesmo sentido de “obrigatório” em que é obrigatório indicar se estamos falando de leões no singular ou no plural quando se fala de leões em inglês. Descrever um teste experimental idealizado para essa hipótese não é difícil. Deixe-nos introduzir uma pessoa a quem podemos chamar de linguista *confederado*. O confederado será (ou pretende ser) um adulto cujo domínio do inglês é geralmente excelente, mas que, por alguma razão (cresceu numa cultura alienígena? foi criado em um monastério?) falhou totalmente na aquisição da palavra “tigre”. O confederado dirá a palavra “tigre” ou, melhor ainda, apontará para ela (como se não tivesse certeza de como pronunciá-la), e perguntará “o que esta palavra significa?” ou “o que é isto?” ou algo assim. Ignorando todas as coisas que dão errado com os experimentos na prática, o que nossa hipótese implica é que informantes deveriam tipicamente dizer ao confederado que tigres são, *inter alia*, listrados. Em vez de confiar nos confederados, pode-se esperar que o linguista estudasse o aprendizado do inglês por uma criança. Mas crianças aprendendo sua língua nativa não são ensinadas tanto quanto os filósofos supõem; elas aprendem, mas não são ensinadas, como Chomsky enfatizou. Ainda assim, as crianças por vezes, fazem perguntas como “o que é um tigre?” e nossa hipótese implica que nesses casos também os informantes deveriam dizer a elas, *inter alia*, que tigres são listrados. Mas um problema é que os informantes são provavelmente como nossos pais, e deve-se considerar os caprichos da interação entre pais e filhos, bem como seu temperamento e a atenção despendida.

Seria fácil especificar um grande número de implicações “operacionais” adicionais de nossa hipótese, mas fazer isso não teria nenhum valor particular. O fato é que somos falantes plenamente competentes do inglês, com um bom senso infernal de quais são nossas obrigações linguísticas. Fingir que estejamos na posição de marcianos com respeito ao inglês, não é uma via para a clareza metodológica. Afinal, foi apenas quando a abordagem operacional foi abandonada que a linguística transformacional floresceu e transformou-se em uma bela ciência.

Desse modo, se alguém me perguntasse pelo significado de “tigre”, sei perfeitamente bem o que diria a ele. Diria-lhe que tigres são felinos, diria algo sobre seu tamanho, que são amarelos com listras pretas, que (por vezes) vivem na selva e são ferozes. Poderia também dizer outras coisas, dependendo do contexto e da razão de seu questionamento; mas os itens acima, salvo possivelmente a parte sobre a selva, eu consideraria como *obrigatórios* transmitir. Não tenho de fazer uma experiência para saber que isso é o que considero obrigatório transmitir, e estou certo de que, aproximadamente, isso também é o que outros falantes consideram como obrigatório transmitir. É claro que há algumas variações de um idioleto a outro; a característica de ter listras (afora as relações entre a figura e a base, por exemplo:

têm listras pretas sobre uma base amarela, que é o modo como eu os vejo, ou seriam listras amarelas sobre uma base preta?) seria encontrada em todos os idioletos normais, mas alguns falantes poderiam considerar a informação de que tigres (estereotipicamente) não habitam selvas como sendo obrigatório, ao contrário de outros falantes. Alternativamente, alguns aspectos do estereótipo (ser um grande felino, listras) poderiam ser considerados como obrigatórios, e outros como *opcionais*, a partir do modelo de certos aspectos sintáticos. Mas não perseguiremos essa possibilidade aqui.

## Os “Dois dogmas” de Quine revisitados

Em “Dois dogmas do empirismo”, Quine lançou um ataque poderoso e saudável à distinção analítico-sintético, que está na moda atualmente. A distinção tinha crescido e se tornado uma verdadeira devoradora filosófica de homens: analítico *igual* a necessário *igual* a não-revisável em princípio *igual* a qualquer verdade que o filósofo individual queira dispensar. Mas o ataque de Quine por si só foi longe demais em certos aspectos; algumas classes limitadas de frases analíticas podem ser salvas, nós achamos.<sup>24</sup> Mais importante, o ataque foi posteriormente interpretado, pelo próprio Quine e outros, como comprometendo a noção inteira de significado a partir da queda da distinção analítico-sintético. Embora tenhamos deixado claro que concordamos que a noção tradicional de significado tem sérios problemas, nosso projeto neste ensaio é construtivo, não destrutivo. Revisamos a noção de significado, não a sepultamos. Assim, será útil ver como os argumentos de Quine saem-se contra nossa revisão.

Os argumentos de Quine contra a noção de analiticidade podem ser reduzidos basicamente ao que se segue: nenhuma importância comportamental pode ser ligada à noção. Seu argumento (novamente simplificando um pouco) foi que havia, basicamente, apenas dois candidatos a indicador comportamental de analiticidade, e ambos são totalmente insatisfatórios, mas por razões diferentes. O primeiro indicador comportamental é a *centralidade*: muitos filósofos contemporâneos chamam uma frase de analítica se, uma comunidade (digamos, os senhores de Oxford) efetivamente sustenta que ela é imune à revisão. Mas, Quine argumenta persuasivamente, a imunidade máxima à revisão não é uma prerrogativa exclusiva de frases analíticas. As frases que expressam as leis fundamentais da física (por exemplo, a conservação da energia) podem muito bem gozar de imunidade comportamental máxima à revisão, ainda que dificilmente seja costumeiro ou plausível classificá-las como analíticas. Quine, contudo, não conta com a mera implausibilidade de classificar como analíticos todos os enunciados que relutamos em abandonar. Ele indica que “imunidade à revisão” é, na história real da ciência, uma *questão de grau*. Na prática efetiva da ciência racional, não há tal coisa como a imunidade *absoluta* à revisão. Assim, identificar analiticidade com imunidade à revisão alteraria a noção de dois modos fundamentais: analiticidade tornar-se-ia uma questão de grau, e não haveria nada que fosse uma frase absolutamente analítica. Isso se afasta tanto da noção de analiticidade clássica de Carnap-Ayer-*et al.* que Quine acha que se *isso* é aquilo de que queremos falar, então seria menos enganoso introduzir um termo diferente, digamos, o de *centralidade*.

O segundo indicador comportamental é *ser chamado de* “analítico”. Com efeito, alguns filósofos tomam a marca característica da analiticidade como sendo a

<sup>24</sup> Ver “The analytic and synthetic” (in: H. Putnam, *Mind, language and reality: Philosophical papers*, vol. 2. Cap. 2. Cambridge: Cambridge University Press, 1975).



de que os informantes treinados (digamos, os senhores de Oxford) *chamam* a frase de analítica. Variantes desse indicador são: que a frase seja dedutível das frases de uma lista finita no topo da qual alguém que esteja na relação ser estudante de pós-graduação com Carnap imprimiu as palavras “Postulado de Significação”; que a frase seja obtível a partir de um teorema lógico substituindo-se sinônimos por sinônimos. A última dessas variantes parece promissora, mas Quine lançou contra ela a questão “qual é o critério de sinonímia?”. Um critério possível poderia ser que as palavras  $W_1$  e  $W_2$  são sinônimas se, e somente se, o bicondicional ( $x$ ) ( $x$  está na extensão de  $W_1 \equiv x$  está na extensão de  $W_2$ ) for *analítico*; mas isso nos leva de volta em um círculo. Outra possibilidade seria que as palavras  $W_1$  e  $W_2$  são sinônimas se, e somente se, informantes treinados as chamam de sinônimas; mas esse é apenas o nosso segundo indicador numa forma levemente revisada. Uma linha promissora é a que as palavras  $W_1$  e  $W_2$  são sinônimas se, e somente se,  $W_1$  e  $W_2$  forem intercambiáveis (isto é, se as palavras podem ser trocadas) *salva veritate* em todos os contextos de uma classe adequada. Mas Quine convincentemente mostrou que essa proposta também nos leva a um círculo. Assim, o segundo indicador reduz-se a isto: uma frase é analítica se ela ou alguma expressão, ou sequência de pares ordenados de expressões, ou conjunto de expressões relacionadas à frase de certos modos específicos está em uma classe a cujos membros os informantes treinados aplicam certos *sons*: ou o *som* ANALÍTICO, ou o *som* POSTULADO DE SIGNIFICAÇÃO, ou o *som* SINÔNIMAS. Em última análise, essa proposta deixa “analítico” etc., como *sons inexplicáveis*.

Embora Quine não discuta isso explicitamente, está claro que tomar a interseção de dois indicadores comportamentais insatisfatórios não seria mais satisfatório. Explicar a analiticidade de uma frase como consistindo em centralidade *mais* ser chamada de ANALÍTICA, é apenas dizer que as frases analíticas são uma subclasse das frases centrais sem de modo algum dizer em que consiste a excepcionalidade da subclasse. Com efeito, a conclusão de Quine é que a analiticidade é ou a centralidade mal concebida ou não é nada.

Apesar do argumento vigoroso de Quine, muitos filósofos continuam abusando da noção de analiticidade, frequentemente confundindo-a com um suposto grau máximo de centralidade. Confrontados com as alternativas de Quine, eles elegeram identificar analiticidade com centralidade, e pagam o preço – o preço de classificar frases que obviamente parecem sintéticas, tal como “o espaço tem três dimensões” como sendo analíticas, e o preço de tentar manter a concepção de que há, afinal de contas, algo absolutamente não-revisável na ciência, apesar da evidência impressionante em contrário. Mas essa linha pode ser destruída atrelando os argumentos de Quine com um argumento importante de Reichenbach.

Reichenbach<sup>25</sup> mostrou que existe um *conjunto* de princípios, cada um dos quais Kant teria considerado como sintético *a priori*, mas cuja conjunção é incompatível com os princípios da relatividade especial e da covariância geral. (Esses incluem a indução normal, a continuidade do espaço e o caráter Euclidiano do espaço.)

Um kantiano pode coerentemente continuar sustentando a geometria euclidiana, aconteça o que acontecer; mas em seguida a experiência pode forçá-lo a abandonar a indução normal ou a continuidade do espaço. Ou ele pode continuar sustentando a indução normal e a continuidade do espaço, venha o que vier; mas então a experiência pode forçá-lo a abandonar a geometria Euclidiana (isso

<sup>25</sup> H. Reichenbach, *The theory of relativity and a priori knowledge* (Los Angeles: University of California Press 1965), p. 31.

acontece no caso em que o espaço físico não é sequer homeomórfico a qualquer espaço Euclidiano). Em seu artigo no volume Schilpp (1951)<sup>26</sup>, Reichenbach fornece essencialmente o mesmo argumento sob uma forma ligeiramente diferente.

Aplicado ao nosso contexto presente, o que isso mostra é que há princípios que os filósofos que gostam da noção exagerada de analiticidade, e em particular, filósofos que identificam analiticidade com não-revisabilidade (máxima), classificariam como sendo analíticos, mas cuja conjunção tem consequências empíricas testáveis. Assim, ou a identificação de analiticidade com centralidade tem de ser abandonada de uma vez por todas, ou temos que abandonar a ideia de que analiticidade é fechada sob a conjunção, ou temos de engolir a consequência infeliz de que uma frase analítica pode ter consequências empíricas testáveis (e desse modo que uma frase *analítica* pode ser *empiricamente falsa*).

Não é acidental, a propósito, que as frases que Kant teria classificado como sintéticas *a priori* seriam classificadas por esses empiristas tardios como analíticas. Seu propósito em inchar a noção de analiticidade foi precisamente dissolver o problema de Kant identificando *aprioridade* com analiticidade, e então identificar analiticidade, por sua vez, com verdade por convenção. (Esse último passo também foi criticado devastadoramente por Quine, mas sua discussão nos afastaria de nosso tema).

Outros filósofos tentaram responder a Quine distinguindo entre *frases e enunciados* [*statements*]: todas as frases são revisáveis, eles concordam, mas alguns *enunciados* não são. Revisar uma frase não é mudar de opinião sobre o enunciado anteriormente expresso por aquela frase apenas no caso de a frase (quer dizer, o objeto sintático mais o seu significado) depois da revisão não ser, de fato, sinônima à frase anterior à revisão, isto é, apenas no caso de a revisão ser um caso de mudança de significado e não de mudança de teoria. Mas, (1) isto se reduz imediatamente à proposta de explicar a analiticidade em termos de sinonímia; e (2) se há algo que Quine contribuiu decisivamente para a filosofia, esse algo foi a compreensão de que mudança de significado e mudança de teoria não podem ser separadas de maneira nítida. Não concordamos com Quine que mudança de significado não possa ser definida de modo algum, mas disso não segue que a dicotomia “mudança de significado *ou* mudança de teoria” seja defensável. Descobrir que vivemos em um mundo não-euclidiano *pode* mudar o significado de “linha reta” (isso aconteceria no caso – um tanto improvável – de que algo como o postulado das paralelas fosse parte do estereótipo de ser uma reta); mas isso não seria uma *mera* mudança de significado. Em particular, isso não seria uma mudança da *extensão*: assim, não seria correto dizer que o postulado das paralelas era “verdadeiro no sentido que as palavras tinham antes”. Do fato de que abandonar uma frase *S* envolveria uma mudança de significado, não se segue que *S* é *verdadeira*. Os significados podem não se encaixar no mundo; e a mudança de significado pode ser forçada por descobertas empíricas.

Embora não estejamos, neste artigo, tentando explicar uma noção de analiticidade, estamos tentando explicar uma noção que pode parecer intimamente relacionada, a noção de significado. Assim, pode parecer que os argumentos de Quine também iriam contra a nossa tentativa. Vamos verificar isso então.

Em nossa concepção, há um sentido perfeitamente aceitável em que ser listrado é parte do significado de “tigre”. Mas não se segue, em nossa concepção, que “tigres são listrados” seja analítica. Se uma mutação ocorresse, todos os tigres

<sup>26</sup> P. Schilpp (ed.), *Albert Einstein philosopher-scientist* (Chicago: Open Court, 1951).

poderiam tornar-se albinos. A comunicação pressupõe que eu tenha um estereótipo de tigres que inclui listras, e que você tenha o estereótipo de tigres que inclui listras, e que eu saiba que o seu estereótipo inclui listras, e que você saiba que o meu estereótipo inclui listras, e que você sabe que eu sei... (e assim por diante, à la Grice, para sempre). Mas isso não pressupõe que qualquer estereótipo particular seja *correto*, ou que a maioria de nossos estereótipos permanecerá para sempre correta. A obrigatoriedade linguística não pretende ser um indicador da não-revisabilidade ou mesmo da verdade. Assim, podemos sustentar que “tigres são listrados” é parte do significado de “tigre” sem ficarmos presos aos problemas da analiticidade.

Desse modo, os argumentos de Quine contra identificar analiticidade com centralidade não são argumentos contra identificar um aspecto como sendo “parte do significado” de X com ele ser obrigatoriamente incluído no estereótipo de X. O que dizer do argumento dos “sons” de Quine?

É claro que indícios sobre o que as pessoas *dizem*, incluindo observações metalinguísticas explícitas, são importantes tanto na “semântica” quanto na sintaxe. Assim, se um falante aponta para um *molusco* e pergunta “isto é um tigre?”, as pessoas provavelmente dariam gargalhadas. (Quando elas parassem de rir,) poderiam dizer ele não sabe o significado de “tigre”, ou “ele não sabe o que são tigres”. Tais comentários podem ser úteis ao linguista. Mas não estamos *definindo* o estereótipo em termos de tais comentários. Dizer que ser “como um grande gato” é parte do significado de tigre, não é meramente dizer que a aplicação de “tigre” a algo que não é “como um grande gato” (e tampouco um tigre) provocaria certos *sons*. Vale dizer, os falantes adquirem a informação de que “tigres são (estereotipicamente) como grandes gatos” do mesmo modo como adquirem a palavra “tigre” e que sentem uma obrigação de garantir que aqueles a quem ensinam o uso da palavra ajam do mesmo modo. Informações sobre um mínimo de habilidades exigidas para fazer parte da comunidade linguística é uma informação significativa; nenhuma circularidade do tipo criticada por Quine aparece aqui.

### A tradução radical

O que nossa teoria não faz, ao menos por si só, é resolver o problema da “tradução radical” de Quine (isto é, tradução de uma língua/cultura alienígena). Não podemos traduzir nosso hipotético cheroquoi para o inglês correlacionando estereótipos, simplesmente porque descobrir o que o estereótipo de, digamos, *wa'arabi* envolve traduzir proferimentos do cheroquoi. Por outro lado, a restrição segundo a qual cada palavra em cheroquoi deveria corresponder à sua imagem em inglês sob a função-tradução no que diz respeito ao estereótipo (ou corresponder aproximadamente, uma vez que em muitos casos a correspondência exata pode não ser alcançada), coloca uma *restrição* severa à função-tradução. Uma vez que tenhamos conseguido traduzir o vocabulário básico do cheroquoi, podemos começar a elucidar os estereótipos, e isso servirá tanto para restringir futuras traduções quanto para verificar a correção interna da parte da função-tradução já construída.

Mesmo onde podemos determinar estereótipos (relativamente, digamos, a uma tentativa de tradução do “vocabulário básico”), isso não basta em geral para determinar uma tradução única. Desse modo, as palavras alemãs *Ulme* e *Buche* têm o mesmo estereótipo de olmo; mas *Ulme* significa “olmo”, enquanto *Buche* significa “faia”. No caso do alemão, o fato de que *Ulme* e ‘olmo’ são cognatas, poderia indicar

a tradução correta (embora isso esteja longe de ser infalível – em geral, palavras cognatas não são sinônimas); mas no caso do grego não temos nenhum indício desse tipo a respeito de qual das duas palavras *oxya*, *ptelea* significa *olmo* e qual significa *faia*; teríamos de simplesmente encontrar um grego que pudesse distinguir olmos de faias (ou *oxya* de *ptelea*). O que isso ilustra é que podem não ser as disposições *típicas* dos falantes a assentir e dissentir que o linguista tem de procurar descobrir; devido à divisão do trabalho linguístico, é frequentemente necessário para o linguista avaliar quem são os especialistas com respeito a *oxya*, ou *wa'arabi*, ou *gavagai*, ou o que quer que seja, antes que possa tentar adivinhar qual a extensão socialmente determinada de uma palavra. Então, essa extensão socialmente determinada e o estereótipo do falante *típico*, ainda que não seja um especialista, funcionarão *ambas* como restrições sobre a função-tradução. Descobrir que o estereótipo de *oxya* é radicalmente diferente do estereótipo de “olmo”, desqualificaria a tradução de *oxya* por “olmo” em todos os contextos exceto os mais extensionais. Mas a descoberta de que a *extensão* de *oxya* não é nem mesmo aproximadamente a classe de olmos, eliminaria completamente essa tradução em todos os contextos.

Notar-se-á que já ampliamos a totalidade dos fatos que contam como evidência para uma função-tradução para além da base austera que Quine permite em *Word and object*. Por exemplo, o fato de que falantes digam tal e tal quando o linguista “confederado” aponta para a palavra *oxya* e pergunta “o que isto quer dizer?” ou “o que é isto?” ou o que quer que seja, não é permitido por Quine (como algo que o linguista possa “saber”) com base em que esse tipo de “conhecimento” pressupõe já se ter traduzido a questão “o que essa palavra significa?”. Todavia, se Quine está disposto a supor que é possível *de alguma maneira* adivinhar quais palavras significam *assentimento* ou *dissentimento* na linguagem alienígena, não parece de forma alguma insensato supor que se possa de alguma maneira transmitir a um falante nativo que não se compreende uma palavra. Não é necessário que se descubra uma locução na língua alienígena que literalmente signifique “o que essa palavra quer dizer?” (por oposição a: “não compreendo essa palavra”, ou “essa palavra não me é familiar”, ou “estou intrigado por essa palavra” etc.). Talvez apenas dizendo a palavra *oxya*, ou o que for, com um tom de perplexidade, seja suficiente. Por que a *perplexidade* deveria ser menos acessível ao linguista do que o *assentimento*?

Além disso, estamos aproveitando-nos do fato de que a segmentação em *palavras* revelou-se linguisticamente universal (e até mesmo existem testes para a segmentação de palavras e morfemas que são independentes do significado). Claramente, não há razão fundamentada para permitir que o linguista profira frases completas e busque assentimento ou dissentimento, ao mesmo em que não permitimos que ele profira palavras e morfemas em tom de perplexidade.

Repito, não estou propondo que ampliar o conjunto de evidências dessa maneira resolve o problema da tradução radical. O que ela faz é adicionar restrições adicionais à classe de candidatos admissíveis a uma tradução correta. O que acredito é que ampliando a classe de restrições, pode-se determinar uma tradução única, ou uma tradução tão única quanto somos capazes de obter na prática. Mas as restrições que vão além da teoria linguística propriamente dita terão de ser usadas, em minha opinião. Também haverá restrições sobre que tipos de crenças (e conexões entre crenças, e conexões de crenças com a cultura e o mundo) que podemos razoavelmente imputar às pessoas. A discussão desses assuntos será adiada para um outro artigo.

## Uma crítica da teoria semântica davidsoniana

Em uma série de publicações, Donald Davidson apresentou a sugestão interessante de que uma teoria semântica de uma linguagem natural pode ser modelada naquilo que os lógicos matemáticos chamam de *definição de verdade* para uma linguagem formalizada. Despida de technicalidades, o que essa sugestão diz basicamente é que se pode ter um conjunto de regras especificando (1) para cada palavra, sob quais condições essa palavra é verdadeira de algo (para palavras em que o conceito de uma extensão faz sentido; todas as outras palavras devem ser tratadas como sincategoremáticas); (2) para frases com mais de uma única palavra, uma regra é fornecida especificando as condições sob as quais a frase é verdadeira como uma função da maneira como é construída, a partir de frases menores (contando palavras como se fossem frases de uma palavra só, por exemplo “neve” como sendo “isto é neve”). A escolha de frases de uma palavra só como sendo o ponto de partida é a minha interpretação do que Davidson pretende. De qualquer modo, ele pede que se comece com um conjunto *finito* de frases *curtas* para as quais as condições de verdade devem ser estabelecidas *diretamente*. A intenção de (2) não é que deveria haver uma regra para cada frase não tratada em (1), uma vez que isso exigiria um número infinito de regras, mas que deveria haver uma regra para cada *tipo* de frase. Por exemplo, numa linguagem formalizada, uma das regras do tipo (2) poderia ser: se  $S$  é  $(S_1 \ \& \ S_2)$  para quaisquer frases  $S_1$ ,  $S_2$ , então  $S$  é verdadeiro se e somente se  $S_1$ ,  $S_2$  são ambas verdadeiras.

Note-se que, no exemplo recém dado, a condição de verdade especificada para frases do tipo  $(S_1 \ \& \ S_2)$  realiza a tarefa de especificar o significado de “&”. Mais precisamente, ela especifica o significado da estrutura  $(— \ \& \ —)$ . Esse é o sentido em que uma definição de verdade pode ser uma teoria do significado. A alegação de Davidson é que a teoria *inteira* do significado para uma linguagem natural pode ser dada dessa forma. Não há dúvida de que regras do tipo ilustrado podem dar o significado de algumas palavras e estruturas. A questão é: que razão há para se pensar que o significado da maioria das palavras possa ser dado desse modo, isso sem falar no significado de todas as palavras?

A dificuldade óbvia é esta: para muitas palavras, uma definição de verdade extensionalmente correta pode ser dada que não é em nenhum sentido uma teoria do significado da palavra. Por exemplo, considere “*Água*” é verdadeiro de  $x$  se, e somente se,  $x$  é  $H_2O$ . Essa é uma definição de verdade extensionalmente correta de “água” (falando estritamente, não é uma definição de verdade, mas uma definição de “verdadeiro de” – isto é, uma definição de *satisfação-no-sentido-de-Tarski*, mas não iremos nos preocupar com tais sutilezas aqui). Ao menos ela é extensionalmente correta, se ignorarmos o problema de que água com impurezas é também chamada de “água” etc. Agora suponha que a maioria dos falantes não *sabe* que água é  $H_2O$ . Então essa fórmula de modo algum nos diz algo acerca do *significado* de “água”. Pode ser interessante a um químico, mas não conta como uma teoria do significado do termo “água”. Ou conta como uma teoria da *extensão* do termo “água”, mas Davidson nos prometeu mais do que apenas isso.

Davidson está bem consciente dessa dificuldade. Sua resposta (em conversas particulares, ao menos) é que precisamos desenvolver uma teoria da *tradução*. Isso ele, assim como Quine, considera ser o problema real. Relativizada a tal teoria (relativizada ao que admitidamente ainda não temos), a teoria é basicamente esta: queremos um sistema de definições de verdade que seja

simultaneamente um sistema de traduções (ou traduções aproximadas, se uma tradução perfeita não for alcançável). Se tivéssemos uma teoria que especificasse o que vem a ser uma boa tradução, então poderíamos excluir a definição de verdade acima mencionada para a “água” como desinteressante, com base em que  $x$  é  $H_2O$  não é uma tradução aceitável ou mesmo uma tradução aproximada de  $x$  é *água* (em uma comunidade pré-científica), mesmo se  $\text{água} = H_2O$  seja verdadeira.

Isso se aproxima perigosamente de dizer que uma teoria do significado é uma definição de verdade mais uma teoria do significado. (Se tivéssemos presunto e ovos, teríamos presuntos e ovos – se tivéssemos presunto e se tivéssemos ovos.) Mas essa história padece de algo pior do que ser uma mera promissória, como veremos.

Uma segunda alegação de Davidson é que a teoria da tradução que ainda não temos é necessariamente uma teoria cujas unidades básicas são *frases* e não *palavras*, pois os *indícios* que temos em linguística necessariamente consistem de assentimento e dissentimento a frases. Palavras podem ser manipuladas, Davidson argumenta, tratando-as como frases (“água” como “isto é água” etc.).

Como é que este ambicioso projeto de construir uma teoria do significado na forma de uma definição de verdade restringida por uma teoria da tradução testada pela “a única evidência que temos”, que são as disposições dos falantes para usar as frases, sai-se de acordo com a concepção que estamos aqui apresentando?

Nossa resposta é que a teoria não pode obter êxito em princípio. Em casos especiais, tais como a palavra “e” no seu sentido veritativo-funcional, uma definição de verdade (estritamente falando, uma cláusula daquilo que os lógicos chamam de uma “definição de verdade” – a soma total de todas as cláusulas é a definição indutiva de “verdade” para a linguagem particular) pode dar o significado da palavra ou da estrutura porque o estereótipo associado à palavra (se queremos falar de um estereótipo no caso de uma palavra como “e”) é tão forte que na realidade constitui uma condição necessária e suficiente. Se todas as palavras fossem como “e” e “solteiro”, o programa poderia ter êxito. E Davidson certamente deu uma contribuição importante salientando que a linguística tem de lidar com condições de verdade indutivamente especificadas. Mas na grande maioria das palavras, os requisitos de uma teoria da verdade e os requisitos de uma teoria do significado são mutuamente incompatíveis, ao menos no caso inglês-inglês. Mas o caso inglês-inglês – o caso em que tentamos fornecer uma teoria significativa do significado de palavras do inglês que seja ela mesma expressa em inglês – certamente é o caso básico.

O problema é que em geral as únicas expressões que são coextensivas à  $X$  e têm aproximadamente o mesmo estereótipo que  $X$ , são expressões que contêm o próprio  $X$ . Se excluirmos tais definições de verdade (estritamente falando, cláusulas, mas vou continuar usando “definição de verdade” tanto para cláusulas individuais quanto para todo o sistema de cláusulas, por simplicidade) como

*“X é água” é verdadeiro se, e somente se, X é água*

com base em que não dizem nada sobre o significado da palavra “água”, e excluirmos definições de verdade tais como

*“X é água” é verdadeiro se, e somente se, X é  $H_2O$*

com base em que o que dizem está errado como uma descrição do *significado* da palavra 'água', então ficaremos com nada.

O problema é que queremos que

*W é verdadeiro de x se e somente se* —

satisfaça as seguintes condições: (1) que a cláusula seja extensionalmente correta (onde — deve ser pensado como uma condição contendo “x”, por exemplo, “x é H<sub>2</sub>O”); (2) que — seja uma *tradução* de *W* – na nossa teoria, isso significaria que o estereótipo associado a *W* é aproximadamente o mesmo que o estereótipo associado a —; (3) que — não contenha *W*, nem variantes sintáticas de *W*. Se tomarmos *W* como sendo, por exemplo, a palavra “olmo”, então não há absolutamente nenhuma maneira de satisfazer todas às três condições simultaneamente. Qualquer condição da forma acima que não contenha “olmo” e que seja extensionalmente correta, conterà um — que é absolutamente terrível como uma *tradução* de “olmo”.

Mesmo onde a linguagem contém dois sinônimos exatos, a situação não é muito melhor. Assim,

*“Urze” é verdadeiro de x se, e somente se, x é tojo*

é verdadeiro, bem como

*“Tojo” é verdadeiro de x se, e somente se, x é urze*

– isto é uma *teoria* do significado de “tojo” e “urze”?

Note-se que a condição (3) é precisamente o que os lógicos *não* impõem às suas definições de verdade.

*“A neve é branca” é verdadeira, se, e somente se, a neve é branca*

é o paradigma de uma definição de verdade no sentido dado pelos lógicos. Mas os lógicos estão tentando fornecer a extensão de ‘verdadeiro’ em relação a uma linguagem particular, não o significado de “a neve é branca”. Tarski teria ido ao ponto de afirmar que estava fornecendo o *significado* (e não apenas a extensão) de “verdadeiro”; mas nunca afirmou que estava dizendo *qualquer* coisa sobre o significado de “a neve é branca”.

Pode ser que o que Davidson realmente pense é que uma teoria do significado, em qualquer sentido sério do termo, seja impossível, e que tudo o que é possível é construir funções de tradução. Se esse é o caso, ele poderia muito bem pensar que a única “teoria do significado” possível para o inglês é aquela que diz que “‘olmo’ é verdadeiro de x se, e somente se, x for um olmo”, “‘água’ é verdadeiro de x se, e somente se, x é água” etc., e apenas raramente esclarecedor como “S<sub>1</sub> & S<sub>2</sub> é verdadeiro se, e somente se, S<sub>1</sub>, S<sub>2</sub> são ambas verdadeiras”. Mas se a “teoria” de Davidson é apenas ceticismo quineano sob o disfarce de uma contribuição positiva ao estudo do significado, então se trata de uma pílula difícil de engolir.

A alegação de que a única evidência disponível ao linguista são as disposições dos falantes com respeito a frases inteiras é, além disso, vazia sob uma interpretação, e simplesmente falsa na interpretação em que ela não é vazia. Se

disposições para dizer certas coisas *quando perguntadas sobre palavras individuais ou morfemas ou estruturas sintáticas* são incluídas na noção de disposições de uso de frases, então a restrição a disposições de uso de frases parece não excluir coisa alguma. Na interpretação não vazia, o que Davidson está dizendo é que o linguista não pode ter acesso a dados como aqueles que os informantes (incluindo o próprio linguista) dizem quando perguntados sobre o significado de uma palavra ou morfema ou estrutura sintática. Nenhuma razão jamais foi dada sobre por que o linguista não poderia ter acesso a tais dados, e é óbvio que os linguistas atuais confiam muito no testemunho dos informantes sobre tais assuntos, no caso de uma língua estrangeira, e em suas próprias intuições como falantes nativos, quando estão estudando suas línguas nativas. Em particular, quando estamos tentando traduzir uma frase inteira, não há razão por que não deveríamos nos guiar pelo nosso conhecimento das propriedades semânticas e sintáticas dos constituintes daquela frase, incluindo a estrutura profunda. Como vimos, há procedimentos para se obter informações sobre constituintes individuais. Vale ressaltar que o procedimento que Quine e Davidson alegam ser o único *possível* – indo de frases inteiras às palavras individuais – é o *oposto* do procedimento sobre o qual todo o sucesso já alcançado no estudo de linguagem natural tem se baseado.

### Crítica da semântica californiana

Quero agora considerar uma abordagem da teoria semântica desbravada pelo Rudolph Carnap tardio. Uma vez que não quero me envolver em questões textuais, não atribuirei a forma particular da concepção que descreverei a nenhum filósofo em particular, mas simplesmente me referirei a ela como “semântica californiana”.

Suporemos a noção de um *mundo possível*. Seja  $f$  uma função definida no “espaço” de todos os mundos possíveis cujo valor  $f(x)$  em qualquer mundo possível  $x$  é sempre um subconjunto do conjunto de entidades em  $x$ . Então  $f$  é dita uma *intensão*. Um termo  $T$  tem significado para um falante  $X$  se  $X$  associa  $T$  com uma intensão  $f_T$ . O termo  $T$  é *verdadeiro* de uma entidade  $e$  em um mundo possível  $x$  se, e somente se,  $e$  pertence ao conjunto  $f(x)$ . Em vez de usar o termo ‘associado’, o próprio Carnap tendeu a falar da “apreensão” de intensões; mas, claramente, o que se estava pretendendo era não apenas que  $X$  “apreende” a intensão  $f$ , mas que ele apreende *que*  $f$  é a intensão de  $T$  – isto é, que ele *associe*  $f$  com  $T$  de algum modo.

Claramente, essa imagem do que vem a ser compreender um termo diverge da história narrada neste artigo. A resposta de um semântico californiano seria que a semântica californiana é uma descrição de uma linguagem *ideal*; que a linguagem real é *vaga*. Em outras palavras, um termo  $T$  na linguagem real não tem uma intensão precisa única; ela tem um conjunto – possivelmente um conjunto *fuzzy* – de intensões. Não obstante, o primeiro passo em direção à descrição da linguagem natural é certamente estudar a idealização em que cada termo  $T$  tem exatamente uma intensão.

(Em seu livro *Meaning and necessity*, Carnap emprega uma formulação superficialmente diferente: uma intensão é simplesmente uma *propriedade*. Uma entidade  $e$  pertence à extensão de um termo  $T$  somente no caso de  $e$  ter qualquer propriedade que esteja na intensão de  $T$ . A última formulação em termos de funções  $f$ , como descritas acima, evita tomar a noção de *propriedade* como primitiva.)

A primeira dificuldade com essa posição é o uso da noção totalmente inexplicada de *apreender* uma intensão (ou, em nossa reformulação da posição, *associar* uma intensão a um termo). Identificando intensões com entidades  $f$  da



teoria dos conjuntos, fornece uma compreensão “concreta” da noção de intensão no estilo matemático atual (relativo às noções de mundo possível e conjunto), mas ao custo de tornar muito difícil perceber como alguém poderia ter uma intensão em sua mente, ou o que seria pensar sobre uma intensão, ou “apreendê-la” ou “associá-la” ao que quer que seja. Não basta dizer que pensar sobre uma intensão seja usar uma palavra ou um substituto funcional para uma palavra (por exemplo, o análogo de uma palavra no “código cerebral”, se, como parece provável, o cérebro “computa” em um “código” que tem analogias com e possivelmente empréstimos da linguagem; ou uma forma de pensamento tal como uma imagem ou um símbolo privado, em casos onde são empregados em pensamento) que se refere à intensão em questão, uma vez que a *referência* (isto é, estar na extensão de um termo) acaba de ser definida em termos de *intensão*. Embora a caracterização do que seja o pensar de uma entidade abstrata tal como uma função ou propriedade é certamente correto, no presente contexto é patentemente circular. Mas nenhuma caracterização não-circular dessa noção fundamental da teoria jamais foi fornecida.

Essa dificuldade está relacionada a uma dificuldade geral na filosofia da matemática apontada por Paul Benacerraf.<sup>27</sup> Benacerraf observou que as filosofias da matemática tendem a se encaixar sobre uma das seguintes bases: ou dão conta do que são os objetos matemáticos e da necessidade da verdade matemática e deixam de dar conta do fato de que as pessoas podem *aprender* matemática, podem fazer *referência* a objetos matemáticos etc., ou, em vez disso, dão conta destes últimos fatos e deixam de dar conta dos primeiros. A semântica californiana dá conta do que *são* as intensões, mas não fornece nenhuma explicação que não seja completamente circular de como é que podemos “apreendê-las”, associá-las a termos, pensar sobre elas, *referir-se* a elas etc.

Carnap pode não ter notado essa dificuldade por causa de seu verificacionismo. Em seus anos iniciais, Carnap concebeu a compreensão de um termo como a posse da *habilidade de verificar* se uma dada entidade cai ou não na extensão do termo. Em termos de intensões: “apreender” uma intensão, então, seria possuir a habilidade de verificar se uma entidade  $e$  em qualquer mundo possível  $x$  pertence a  $f(x)$  ou não. O Carnap tardio modificou essa concepção, reconhecendo que, como diz Quine, frases se defrontam com o tribunal da experiência coletivamente e não individualmente. Não há tal coisa como o modo de verificar que um termo  $T$  seja verdadeiro de uma entidade, em geral, independente do contexto de um conjunto particular de teorias, hipóteses auxiliares etc. Talvez Carnap mantivesse que algo como a sua teoria anterior estava correta para uma classe limitada de termos, os assim chamados “termos observacionais”. Nossa própria concepção é que a teoria verificacionista do significado é falsa tanto em sua ideia central quanto para os termos observacionais, mas não tentaremos discutir isso aqui. De qualquer modo, se alguém *não* for um verificacionista, então fica difícil ver a semântica californiana como sendo uma teoria, uma vez que a noção de *apreender* uma intensão foi deixada totalmente inexplicada.

Em segundo lugar, se supusermos que “apreender uma intensão” (associar uma intensão a um termo  $T$ ) deveria ser um *estado psicológico* (num sentido estrito), então a semântica californiana está comprometida com os princípios (1) e (2) que criticamos na primeira parte deste artigo. Ela tem de sustentar que o estado psicológico de um falante determina a intensão de seus termos que, por sua vez, determina a extensão de seus termos. Seguir-se-ia que se dois seres humanos

<sup>27</sup> P. Benacerraf, “Mathematical truth”, *Journal of Philosophy*, 70.19 (1973): 661-679.

estão no mesmo estado psicológico, então necessariamente atribuem a mesma extensão a todos os termos que empregam. Como vimos, isso está totalmente errado para a linguagem natural. A razão pela qual isso está errado, como vimos acima, é em parte que a extensão é determinada socialmente, não pela competência individual isoladamente. Desse modo, a semântica californiana está comprometida com tratar a linguagem como algo privado – ignorando totalmente a divisão do trabalho linguístico. A extensão de cada termo é vista por essa escola como totalmente determinada por algo na cabeça do falante individual que fala sozinho. Uma segunda razão pela qual isso está errado, como também vimos, é que a maioria dos termos são *rígidos*. Na semântica californiana, cada termo é tratado como, em efeito, uma *descrição*. O componente *indexical* do significado – o fato de que nossos termos referem-se a coisas que são semelhantes, de determinadas maneiras, a coisas que designamos rigidamente, a *estas* coisas, ao que chamamos de “água”, ou o que quer que seja, *aqui* – é ignorado.

Mas o que dizer da defesa segundo a qual o semântico californiano não está preocupado com a linguagem real, mas com uma idealização em que “ignoramos a vagueza”, e que os termos da linguagem natural podem ser pensados como estando associados a um conjunto de intensões em vez de uma única intensão bem definida?

A resposta é que uma palavra *indexical* não pode ser representada como uma família vaga de palavras não-indexicais. A palavra “eu”, para tomar um caso extremo, é *indexical* mas não *vaga*. “Eu” não é sinônimo de uma *descrição*; tampouco é sinônimo de um conjunto *fuzzy* de descrições. Semelhantemente, se estamos certos, “água” não é sinônimo nem de uma descrição, nem de um conjunto *fuzzy* de descrições (intensões).

Similarmente, uma palavra cuja extensão é fixada socialmente e não individualmente não é a mesma coisa que uma palavra cuja extensão é vagamente fixada individualmente. A razão pela qual minha “apreensão” individual de “olmo” não fixa a extensão de olmo, não é que a palavra seja vaga – se o problema fosse simplesmente de vagueza, então o fato de que meus conceitos não distinguem olmos de faias implicaria que olmos são faias, tal como eu uso o termo, ou, de qualquer modo, casos limítrofes de faias, e que faias são olmos, ou casos limítrofes de olmos. Em vez disso, a razão é que a extensão de “olmo” no meu dialeto não é fixada pelo que o falante médio “apreende” ou não “apreende” de modo algum; é fixada pela comunidade, incluindo os especialistas, por meio de um processo cooperativo complexo. Uma linguagem que exemplifica a divisão do trabalho linguístico não pode ser aproximada de modo bem sucedido a uma linguagem que possui termos vagos mas nenhuma divisão do trabalho linguístico. Cooperação não é vagueza.

Mas, alguém poderia retrucar: não se poderia substituir nossa linguagem real por uma linguagem em que (1) os termos são substituídos por termos coextensivos que *não* fossem indexicais (por exemplo, “água” por “H<sub>2</sub>O”, supondo-se que “H<sub>2</sub>O” não seja indexical); e (2) eliminamos assim a divisão do trabalho linguístico tornando cada falante um especialista em todos os tópicos?

Responderemos essa questão negativamente. Mas suponha, por um momento, que a resposta fosse “sim”. Que importância isso teria? A linguagem “ideal” não seria em nenhum sentido similar à nossa linguagem real, e tampouco a diferença seria uma questão de “vagueza da linguagem natural”.

De fato, contudo, não se pode realizar a substituição, pela excelente razão de que *todas* as palavras para espécies naturais e palavras para magnitudes físicas são

indexicais do modo como descrevemos; “hidrogênio”, e consequentemente “H<sub>2</sub>O”, são indexicais tanto quanto “água”. Talvez termos para “dados dos sentidos” não sejam indexicais (afora os termos para si-mesmo), se é que existem. Mas “amarelo” como um predicado de coisas é indexical pela mesma razão que “tigre”. Mesmo se algo *parece* amarelo pode não ser amarelo. E não ajuda nada dizer que coisas que parecem amarelas em circunstâncias normais (à perceptores normais) são amarelas; aqui “normal” tem precisamente a característica que chamamos de indexicalidade. Simplesmente não há nenhuma razão para acreditar que o projeto de reduzir nossa linguagem a uma linguagem não-indexical poderia ser em princípio realizada.

A eliminação da divisão do trabalho linguístico poderia, suponho, “em princípio” ser realizada. Mas, se a divisão do trabalho linguístico é, como conjecturei, um universal linguístico, que interesse haveria na possibilidade da existência de uma linguagem que carece de uma característica constitutiva da linguagem *humana*? Um mundo onde cada um é especialista em todos os tópicos é um mundo em que as leis sociais são quase inimaginavelmente diferentes do que são hoje. Qual a *motivação* para tomar tal mundo e tal linguagem como o modelo para a análise da linguagem *humana*?

Incidentalmente, filósofos que trabalham na tradição da semântica californiana recentemente começaram a modificar o esquema para superar justamente esses defeitos. Assim, foi sugerido que uma intensão poderia ser uma função cujos argumentos não são apenas mundos possíveis, mas, talvez, um mundo possível, um falante, e um contexto não-linguístico de enunciação. Isso permitiria a representação de alguns tipos de indexicalidade e alguns tipos de divisão de trabalho linguístico no modelo. Do modo como David Lewis desenvolve essas ideias, “água”, por exemplo, teria a mesma *intensão* (mesma função) em Terra e em Terra Gêmea, mas uma extensão diferente. (Com efeito, Lewis retém a suposição (1) da discussão na primeira parte deste trabalho e abre mão de (2); nós escolhemos abandonar (1) e manter (2).) Não há razão por que os modelos formais desenvolvidos por Carnap e seus seguidores não deveriam se mostrar valiosos quando assim modificados. Nosso interesse aqui tem estado não na utilidade do formalismo matemático, mas na filosofia da linguagem subjacente às versões anteriores da concepção.

## Marcadores semânticos

Se a abordagem aqui sugerida está correta, então há uma grande quantidade de trabalho científico a ser feito na (1) descoberta de que tipos de itens podem aparecer em estereótipos; e (2) na elaboração de um sistema conveniente para representar estereótipos; etc. Esse trabalho não é algo que possa ser feito pela discussão filosófica, contudo. Em vez disso, trata-se de uma província da linguística e da psicolinguística. Uma ideia que pode, creio eu, ser valiosa, é a ideia de um *marcador semântico*. A ideia vem do trabalho de J. Katz e J. A. Fodor. Nós a modificaremos um pouco aqui.

Considere o estereótipo de “tigre” por um momento. Isso inclui tais características como ser um animal, ser como-um-grande-gato, ter listras pretas sobre uma base amarela (listras amarelas sobre uma base preta?), etc. Ora, há algo muito especial sobre a característica *animal*. Em termos da noção de Quine de *centralidade* ou de *não-revisabilidade*, ela é qualitativamente diferente das outras características listadas. Não é impossível imaginar que tigres possam não ser animais (poderiam ser robôs). Mas desenvolvendo essa ideia, eles teriam de ter sido

sempre robôs. Não queremos contar uma história sobre os tigres sendo *substituídos* por robôs, porque então os robôs não seriam tigres. Ou, se não tivessem sido sempre robôs, têm de ter se *tornado* robôs, o que é ainda mais difícil de imaginar. Se tigres são e sempre foram robôs, esses robôs não podem ser “inteligentes” demais, ou então não poderíamos ter um caso em que tigres não são animais – poderíamos, em vez disso, ter descrito um caso em que alguns robôs são animais. Melhor seria torná-los robôs “dirigidos por outros” – digamos, ter um operador em Marte controlando cada movimento por controle remoto. Expor os detalhes disso, eu repito, é difícil, e é curiosamente difícil pensar nesse caso em princípio, e essa é a razão pela qual é fácil cometer o erro de pensar que é “logicamente impossível” para um tigre *não* ser um animal. Por outro lado, não há nenhuma dificuldade em se imaginar um tigre individual que não seja listrado; ele poderia ser um albino. Tampouco é difícil imaginar um tigre individual que não se parece com um grande gato: ele poderia estar terrivelmente deformado. Podemos até mesmo imaginar todos os membros da espécie perdendo suas listras ou ficando horrivelmente deformados. Mas os tigres deixariam de ser animais? Novamente, uma grande dificuldade!

Note-se que não estamos incorrendo no erro que Quine criticou acertadamente, o de atribuir irrevisibilidade absoluta a frases tais como “tigres são animais”, “tigres não podem transformar-se de animais em outra coisa e continuarem sendo tigres”. De fato, podemos descrever casos remotos em que esses enunciados seriam abandonados. Mas mantemos que é *qualitativamente* mais difícil revisar “todos os tigres são animais” do que “todos os tigres tem listras” – de fato, este último enunciado sequer é verdadeiro.

Não apenas características como “animal”, “ser vivo”, “artefato”, “dia da semana”, “período de tempo”, vinculam-se com enorme centralidade às palavras “tigre”, “molusco”, “cadeira”, “terça-feira”, “hora”, mas também fazem parte de um *sistema de classificação* importante e amplamente usado. A centralidade garante que itens classificados sob essas rubricas virtualmente nunca têm de ser *reclassificados*. Assim, essas rubricas são as que naturalmente usamos como indicadores de categorias em muitos contextos. Parece-me razoável que, assim como na sintaxe usamos marcadores como “substantivo”, “adjetivo”, e mais especificamente, “substantivos concretos”, “verbos que tomam uma pessoa como um sujeito e um objeto abstrato” etc. para classificar as palavras, assim também na semântica esses indicadores de categorias deveriam ser usados como marcadores.

É interessante que, quando Katz e Fodor introduziram originalmente a ideia de um marcador semântico, não pretenderam esgotar o significado – aquilo que chamamos de estereótipo – por uma lista de marcadores desse tipo. Em vez disso, os marcadores foram restritos a apenas indicadores de categorias de alta centralidade, que é o que propomos. As características remanescentes foram simplesmente listadas como um “diferenciador”. O esquema deles não é facilmente comparável ao nosso, porque eles queriam que os marcadores semânticos *mais* o diferenciador sempre dessem uma condição necessária e suficiente para o pertencimento na extensão de um termo. Uma vez que tudo isso – marcadores e diferenciadores – deveriam representar o que cada falante sabe implicitamente, eles estavam comprometidos com a ideia de que cada falante conhece implicitamente uma condição necessária e suficiente para o pertencimento na extensão de “ouro”, “alumínio”, “olmo” – o que, como indicamos, não é o caso. Posteriormente, Katz foi mais além e exigiu que *todas* as características constituem-se uma condição *analiticamente* necessária e suficiente para o pertencimento à extensão. Nesse

ponto, abandonou a distinção entre marcadores e diferenciadores. Se todas as características têm, por assim dizer, o grau infinito de centralidade, por que chamar algumas de “marcadores” e algumas de “diferenciais”? Do nosso ponto de vista, sua distinção original entre “marcadores” e “diferenciadores” estava correta – contanto que se abandone a ideia de que o diferenciador fornece (junto com os marcadores) uma condição necessária e suficiente, e a ideia de que qualquer parte disso seja uma teoria da *analiticidade*. Sugerimos que a ideia de marcador semântico é uma contribuição importante, quando tomada como aqui sugerida.

## O significado de significado

Podemos agora resumir o que foi dito sob a forma de uma proposta acerca de como se pode reconstruir a noção de “significado”. Nossa proposta não é a única que poderia ser defendida com base nessas ideias, mas pode servir para encapsular alguns dos pontos principais. Além disso, sinto que isso recupera tanto os usos comuns da fala do senso comum e os da linguística quanto é provável que se seja capaz de preservar convenientemente. Uma vez que, em minha concepção, algo como as suposições (I) e (II) listadas na primeira parte deste artigo estão profundamente incorporadas aos significados da fala comum, e essas suposições são conjuntamente inconsistentes com os fatos, nenhuma reconstrução estará livre de algumas consequências contraintuitivas.

Em suma, minha proposta é definir “significado” sem selecionar um objeto que será identificado com o significado (embora isso possa ser feito no estilo usual da teoria dos conjuntos, se alguém insistir), mas especificando uma forma normal (ou, melhor, um *tipo* de forma normal) para a descrição do significado. Se sabemos o que uma “descrição em forma normal” do significado de uma palavra deveria ser, então, no que me diz respeito, sabemos o que o significado é em qualquer sentido cientificamente interessante.

Minha proposta é que a descrição em forma normal do significado de uma palavra deveria ser uma sequência finita, ou um “vetor”, cujos componentes deveriam incluir certamente o seguinte (talvez seja desejável se dispor de outros tipos de componentes também): (1) os marcadores sintáticos que se aplicam às palavras, por exemplo, “substantivo”; (2) os marcadores semânticos que se aplicam às palavras, por exemplo, “animal”, “período de tempo”; (3) uma descrição das características adicionais do estereótipo, se houver; (4) uma descrição da extensão.

A seguinte convenção é uma parte dessa proposta: os componentes do vetor todos representam uma hipótese sobre a competência de um falante individual, *exceto a extensão*. Dessa forma, a descrição em forma normal de “água” pode ser, em parte:

MARCADORES SINTÁTICOS	MARCADORES SEMÂNTICOS	ESTEREÓTIPO	EXTENSÃO
<i>substantivo de massa, concreto;</i>	<i>tipo natural; líquido;</i>	<i>incolor; transparente; insípido; mata a sede; etc.</i>	<i>H<sub>2</sub>O (com ou sem impurezas)</i>

– isso *não* significa que o conhecimento do fato de que água é H<sub>2</sub>O está sendo imputado ao falante individual ou mesmo à sociedade. Significa que (*nós* dizemos que) a extensão do termo “água” que eles (os falantes em questão) usam é *de fato*

H<sub>2</sub>O. A objeção “quem somos *nós* para dizer o que a extensão do termo *deles* é de fato” foi discutida acima. Note que isso é fundamentalmente uma objeção à noção de *verdade*, e que a extensão é uma parente da verdade e herda seus problemas inerentes.

Chamemos duas descrições de *equivalentes* se forem a mesma, exceto pela descrição da extensão, e as duas descrições forem coextensivas. Então, se o conjunto descrito diversamente nas duas descrições for, *de fato*, a extensão da palavra em questão, e outros componentes na descrição são caracterizações corretas dos vários aspectos de competência que representam, então *ambas* as descrições são corretas. Descrições equivalentes são ambas corretas ou ambas incorretas. Essa é outra maneira de dizer que, embora tenhamos de usar uma *descrição* da extensão para *fornecer* a extensão, pensamos no componente em questão como sendo a *extensão* (o *conjunto*), não a descrição da extensão.

Em particular, a representação das palavras “água” no dialeto terráqueo e “água” no dialeto de Terra Gêmea seria a mesma, exceto que na última coluna, a descrição em forma normal da palavra “água” em Terra Gêmea teria XYZ e não H<sub>2</sub>O. Isso significa, tendo em vista o que acabou de ser dito, que estamos atribuindo a *mesma* competência linguística ao falante típico terráqueo/terráqueo-gêmeo, mas uma extensão diferente à palavra, todavia.

Essa proposta significa que mantemos a suposição (II) de nossa discussão inicial. O significado determina a extensão – por construção, por assim dizer. Mas (I) é abandonada; o estado psicológico de um falante individual não determina o ‘que ele quer dizer’.

Na maioria dos contextos, isso concordará com a maneira como falamos, creio eu. Mas há um paradoxo: suponha que Oscar é um falante bilíngue inglês-alemão. Em nossa concepção, em sua coleção total de dialetos, as palavras “beech” [faia] e *Buche* são *sinônimos exatos*. A descrição em forma normal de seus significados seria idêntica. Mas ele poderia muito bem não saber que são sinônimas! Um falante pode ter dois sinônimos em seu vocabulário e não saber que são sinônimos!

É instrutivo ver como uma falha aparentemente óbvia – “se  $S_1$  e  $S_2$  são sinônimas e Oscar compreende tanto  $S_1$  e  $S_2$ , então Oscar sabe que  $S_1$  e  $S_2$  são sinônimas” – está relacionada à falsidade de (I), em nossa análise. Note-se que se tivéssemos escolhido omitir a extensão como um componente do “vetor-significado”, que é a proposta de David Lewis tal como eu a compreendo, então teríamos o paradoxo de que “olmo” e “faia” teriam o *mesmo significado*, mas extensões diferentes!

Em quase todas as teorias materialistas, crer em uma proposição provavelmente envolve processar alguma *representação* daquela proposição, quer ela seja uma frase em uma linguagem, uma porção do “código cerebral”, uma forma de pensamento, ou o que seja. Os materialistas, e não apenas os materialistas, relutam em pensar que alguém possa acreditar em proposições *pura e simplesmente*. Mas mesmo os materialistas tendem a acreditar que, se alguém crê numa proposição, *qual* representação é empregada é irrelevante. Se  $S_1$  e  $S_2$  são ambas representações que são *acessíveis* a mim, então, se acredito que a proposição expressa por  $S_1$  está sob a representação  $S_1$ , devo também acreditar que está sob a representação  $S_2$  – ao menos, devo assim proceder se tenho qualquer pretensão de racionalidade. Mas, como acabamos de ver, isso não está correto. Oscar pode muito bem crer que *isto* é uma “faia” (há uma placa nela que diz “faia”), mas não crer ou descreer que seja um “*Buche*”. Não se trata apenas de que a

crença seja um processo que envolve representações. Ele crê na proposição (se queremos introduzir 'proposições' de alguma maneira) sob uma representação e não sob outra.

O incrível na teoria do significado é o quanto tempo o assunto esteve sob o domínio de concepções filosóficas equivocadas, e como esses equívocos são fortes. O significado tem sido identificado com uma condição necessária e suficiente por filósofo após filósofo. Na tradição empirista, tem sido identificado com o método de verificação, novamente por filósofo após filósofo. Tampouco esses equívocos têm tido a virtude da exclusividade. Não foram poucos os filósofos que sustentaram que significado = método de verificação = condição necessária e suficiente.

Por outro lado, é incrível como tem sido fraca a força dos fatos. Afinal, o que foi apontado neste ensaio são pouco mais do que verdades prosaicas sobre a maneira como usamos as palavras e o quanto (ou melhor, o quão pouco) efetivamente sabemos quando as empregamos. Minha reflexão pessoal acerca dessas questões começou depois que publiquei um artigo em que confiantemente sustentei que o significado de uma palavra era "uma bateria de regras semânticas"<sup>28</sup>. Em seguida comecei a imaginar como o significado da palavra comum "ouro" poderia ser analisado dessa maneira. E não é que filósofos nunca tivessem considerado tais exemplos: Locke, por exemplo, usa essa palavra como um exemplo e não é incomodado pela ideia de que seu significado seja uma condição necessária e suficiente!

Se há uma razão para que tanto a opinião de leigos quanto a de especialistas tenha se desviado tanto com respeito a um tópico que lida com, afinal de contas, questões que estão na experiência de todo mundo, questões a respeito das quais todos temos mais dados do que sabemos manejar, questões a respeito das quais temos, se abandonamos os preconceitos, intuições bastante claras, então ela tem de estar conectada ao fato de que concepções grotescamente equivocadas da linguagem que são e sempre foram correntes que refletem duas tendências filosóficas específicas e muito centrais: a tendência a tratar a cognição como uma questão puramente *individual* e a tendência a ignorar o *mundo*, na medida em que consiste de mais do que "observações" do indivíduo. Ignorar a divisão do trabalho linguístico é ignorar a dimensão social da cognição. Ignorar o que chamamos de *indexicalidade* da maioria das palavras é ignorar a contribuição do meio circundante. A filosofia da linguagem tradicional, assim como muito da filosofia tradicional, deixa de fora as outras pessoas e o mundo; uma filosofia melhor e uma ciência da linguagem melhor devem conter ambos.

---

<sup>28</sup> Ver "How not to talk about meaning" (in: H. Putnam, *Mind, language and reality: Philosophical papers*, vol. 2. Cap. 6. Cambridge: Cambridge University Press, 1975).